# Design strategy for imaging systems containing freeform surfaces

Dissertation for the acquisition of the academic title

**Doktor-Ingenieur(Dr.-Ing.)**

Submitted to the Council of the Faculty of Physics and Astronomy of the Friedrich-Schiller-Universität Jena

By M.Sc. Chang Liu

Born in Jiangsu, China on 8th February 1991

**Supervisors:**

**Major supervisor:** Prof. Dr. Herbert Gross, Friedrich-Schiller-Universität Jena
**Associate Supervisor:** Prof. Dr. Alois Herkommer, Universität Stuttgart
**Associate Supervisor:** Prof. Dr. Rongguang Liang, The University of Arizona

**Day of the Disputation: 16.05.2019**

# Abstract

Motivations driven by the ever-growing system specifications, alongside the progress in optical manufacturing and metrology open widely the door for implementing freeform surfaces for imaging system performance enhancement. Meanwhile, new challenges regarding handling and analysis are posed for this surface type due to its lack of rotational symmetry. This thesis work focuses on the development of an appropriate aberration analysis tool, as well as implementation strategies of freeform surfaces in imaging systems.

The Aldis theorem, which gives the surface contribution for one ray including all orders, is revisited as the foundation for freeform system analysis. With the original theorem's limitations overcome, this new tool can serve as an alternative for the Seidel diagram, indicating the system sensitivity information for freeform systems.

Furthermore, this thesis investigates the numerical approaches utilized for freeform implementation regarding freeform normalization radius, freeform order, optimization strategy, freeform interaction and early-stage manufacturability assessment for both one-freeform and two-freeform issues. As a consequence, the relation between the choice of the freeform location and system aberration constitutions is revealed with the assistance of the surface eccentricity parameter. A successive implementation strategy for freeform orders and freeform quantity is proposed in the form of a workflow. Diverse system examples are presented in the end as further validation of the findings.

# Zusammenfassung

Die immer stringenter werdenden Spezifikationen sowie die zunehmenden technologischen Entwicklungen in den Herstellungsprozessen erfordern bzw. ermöglichen die Implementierung von Freiformflächen in abbildenden optischen Systemen zur Verbesserung der Abbildungsqualität. Daraus stellen sich neue Herausforderungen hinsichtlich des Handlings und der Analyse dieses Flächentyps aufgrund seiner fehlenden Rotationssymmetrie. Diese Doktorarbeit konzentriert sich auf das Entwickeln eines geeigneten Werkzeuges zur Analyse von Aberrationen sowie von Implementierungsstrategien von Freiformen in optischen Systemen.

Zur Analyse der Freiformsysteme wird Aldis Theorem, welches die Oberflächenbeiträge eines Strahls für sämtliche Ordnungen widergibt, betrachtet und weiterentwickelt. Dadurch dient es als Alternative zum Seidel-Diagramm zur Sensitivitätsanalyse eines Freiformsystems.

Darüber hinaus werden numerische Methoden zur Freiformimplementierung mit Hinblick auf den Freiformnormradius, die Freiformordnung, Optimierungsstrategien, Freiformwechselwirkungen und die Herstellbarkeitsbeurteilung für Systeme mit einer und zwei Freiformen untersucht. Daraus folgend ergibt sich eine Beziehung zwischen der optimalen Freiformposition und den Aberrationsbeiträgen. Diese Beziehung wird mit Hilfe des Flächen-Exzentrizitätsparameters analysiert. Eine schrittweise Implementierungsstrategie der Freiformen in Form eines Workflows wird vorgeschlagen. Diverse Systembeispiele zur weiteren Untermauerung der gewonnenen Erkenntnisse werden präsentiert.

# Contents

# 1 Introduction

Driven by the ever-growing demands on optical system specifications, the employment of non-rotationally symmetric freeform surfaces becomes prevailing alongside the technical breakthroughs in freeform manufacturing and metrology [1, 2]. Unlike traditional rotationally symmetric surfaces, this multiple-degree-of-freedom surface type poses new challenges for optical system design due to the lack of appropriate analysis tools and handling strategies. The understanding and investigation of these aspects are crucial for the better use of freeform surfaces.

Published by Cox [4], introduced first by Aldis, the Aldis theorem delivers the complete surface aberration contribution including all orders for one ray. Similar to the third order Seidel aberration diagram, this theorem can help to locate the sensitive elements of the optical system. Unlike the lower order Seidel aberration, the consideration of the higher orders in the Aldis theorem makes it a useful tool for analysis of more complex systems. Brewer extended the original Aldis theorem to cover aspheres in 1976 [5, 6], but a further generalization of the theorem is still required to include freeform surfaces and more general system geometries. Potentials of the Aldis theorem can be explored to assist freeform system analysis.

Thompson developed the nodal aberration theory based on vectorial wave aberration expansion in 1980 [7, 8]. Analytical expressions up to the $6^{th}$ order are given for all imaging systems with quasi-circular pupils [9-11]. Later, Fuerschbach

investigated further the aberrations generated by specific freeform overlay within the optical system [12-14]. The resulting effects of introducing freeform surfaces can be understood better with his work. The nodal aberration theory is beneficial in identifying the nodal points, as well as in providing the full field aberration display to help to understand the field performance of the system. Nevertheless, the limitation to the 6$^{th}$ order aberration, as well as the complexity of the generated aberrations makes it difficult to be applied analytically in solving multi-surface design problems. A numerical method is inevitable.

In 2005, Yabe developed a numerical method of implementing aspheres in the optical system for performance elevation [15]. The method was later extended to include surface tilts and more generalized surface types [16, 17]. Nonetheless, the relation between the choice of the non-spherical surface location and system aberration constitution is not tackled and detailed strategies for implementing these special types of surfaces are still missing.

The main goal of this thesis is the development and investigation of useful tools for freeform system analysis and optimization. Based on Brewer's generalization of the Aldis theorem, the theorem is extended to include freeform surfaces and to further overcome some limitations of the original theorem. To figure out the proper methods employed in freeform system optimization, numerical approaches utilized for freeform implementation are investigated regarding the freeform normalization radius, the freeform location, the freeform order, the optimization strategy, the potential interaction between two freeforms, and the early-stage manufacturability assessment. Both one-freeform and two-freeform implementation issues are addressed, based on which a workflow is concluded that has the potential to be applied for more general situations. The findings ceased from the investigation are validated on a broader selection of example systems with different system symmetries.

This dissertation is structured as follows. In chapter 2, a brief overview of the state-of-the-art freeform system design methods is provided, followed by the introduction

of nodal aberration theory and Aldis theorem. The freeform surface definition from the surface representation point of view is reviewed and the constraints regarding freeform manufacturability are discussed. Several example optical system types are presented in this chapter, laying the foundation for the following work.

In chapter 3, further extension of the Aldis theorem based on Brewer's generalization is realized. Extensions to freeform surfaces, to an arbitrary image plane, to multiple rays and to more general system geometries are discussed. The possible application of the new set of formulas is deliberated with two examples.

In chapter 4, implementation aspects of one freeform surface and two freeform surfaces are discussed in detail. The link is established between optimal freeform locations and the surface eccentricity and the system aberration constitutions. The optimization strategies in the case of multiple freeform implementations are investigated using two freeform surfaces. Possible interactions between freeforms are discussed. As a guidance for freeform implementation, a workflow based on the presented findings is proposed. Practical realization questions of freeform are considered using the Q-polynomial representation. Considering diamond turning fabrication, the freeform surface sag, the azimuthal gradient, and the acceleration are evaluated, assisting early-stage freeform manufacturability assessment.

More system examples are provided in chapter 5, in which conclusions, drawn from the previous chapter, are validated further with diverse system types.

# 2 Theoretic background

## 2.1 DESIGN PROCESS

The design process of an optical system normally starts with the selection of a start system. Customized merit functions are constructed to optimization the optical system to reach the required specifications. System optimization is nothing but the search for global or local minima on the solution space. Though the strategies involved in optimization determine directly whether the specifications can be reached or not, the choice of the start system has a huge impact on the final performance of the system, if not dominating. For a simple rotationally symmetric system, a start system can be constructed firstly through thin lenses and corrected analytically for third-order aberrations. Nevertheless, as the complexity of the system requirements increases, the selection of the start system is generally done through the searching of existing patents and archive systems, and most of the time depends on the personal experience of optical designers.

As the demands on system features increase, the limitations of the traditional rotationally symmetric systems start to show. Approaches that incorporate freeform surface in start system design are developed. One example is the simultaneous multiple surface (SMS) method [18, 19], which is first used for illumination design and later expanded to imaging system design. The SMS method involves the simultaneous calculation of multiple optical surfaces using multiple bundles of rays

for perfect imaging. Rays coming from multiple fields are coupled into image points by using a limited number of surfaces. The calculated surface profiles are then interpolated to certain freeform representation and optimized according to system specifications. The limitation of this approach is also clear, only a limited number of fields and surfaces (at the moment two surfaces) are allowed.

Another approach is the combination of Gaussian bracket method and nodal aberration theory [20]. In this approach, aberrations of the central field are optimized analytically using spherical surfaces by bringing the nodal points to the center of the aberration field. An initial system with minimum aberrations is achieved with this method, on whose basis freeform surfaces can be implemented. This approach is limited by the complexity of solving the non-linear equations, which requires proper handling of the start values and boundary conditions.

One more approach is being referred to as the conic confocal method. In this method, confocal conic mirrors are rotated around their common focal point for the cancelation of primary aberrations [21, 22]. The surface material and potentially large off-axis use of the aspheres limit this method.

Other than the SMS method, in which the freeform topology is already included in the start system, it is generally more common to implement freeform surfaces later in the design process. The placement of the freeform surfaces becomes a challenging task. S. Mao et al. [23] introduced lens-form parameters as criteria for reducing the number of lenses and the placement of freeforms. In 2005, A. Yabe proposed a method to determine the optimal placement of the aspheric surfaces [15]. Later on, he extended his method by adding surface tilts and allowing for surface types that are more general [16, 17]. This last contribution allows breaking the rotational symmetry of the initial design and creating off-axis systems. However, the relation between freeform locations and system aberration constitution is not tackled, and a detailed strategy for implementing multiple non-spherical surfaces is missing. The progress in optical design resulted in new codes and design tools [24-27], which

provide more correction features to achieve improvement in system performance. However, a systematic approach for optimizing freeform systems is still needed.

## 2.2 PERFORMANCE EVALUATION

For non-rotationally symmetric optical systems, some classical performance measures based on ray tracing still work i.e. spot diagram. However, due to the three-dimensional (3D) geometry of the layout, performance measures like Seidel aberration and Zernike coefficients are of decreased credibility. The extension from the two-dimensional (2D) layout to the 3D layout increases the system complexity drastically and raises new challenges for finding suitable aberration theories.

### 2.2.1 NODAL ABERRATION THEORY

Nodal aberration theory was first introduced by Shack and Thompson in 1980 [1], which was later further developed by Thompson [7-11] and Fuerschbach [12-14]. The theory was mainly used for telescope misalignment analysis in the early phase, and later expanded to non-rotationally symmetric systems without the limitations of small element tilts and decentration. The theory supports analytical wave aberration derivation up to the 6$^{th}$ order for all imaging optical systems with circular pupils, or nearly circular pupils. It is a direct extension of the rotationally symmetric wave aberrations used in optical design.

The conventional wave aberration expansion developed by Hopkins [29] is converted into a vector form [8] using Shack's method. The vectorial wave aberration form is given in Eq. (1), where $\vec{H}$ represents the field position in the image field, and $\vec{\rho}$ represents the pupil position.

$$\begin{aligned} W &= W[(\vec{H}\cdot\vec{H}),(\vec{H}\cdot\vec{\rho}),(\vec{\rho}\cdot\vec{\rho})] \\ &= \sum_{j}\sum_{p}^{\infty}\sum_{n}^{\infty}\sum_{m}^{\infty}(W_{klm})_{j}(\vec{H}\cdot\vec{H})^{p}(\vec{\rho}\cdot\vec{\rho})^{n}(\vec{H}\cdot\vec{\rho})^{m} \end{aligned} \qquad (1)$$

Expansion of Eq. (1) through primary aberrations gives Eq. (2). In this equation, $W_{040}$ is the coefficient for primary spherical aberration, $W_{131}$ is the coefficient for primary coma, $W_{222}$ is the coefficient for primary astigmatism, $W_{220}$ is the coefficient for field curvature, and $W_{311}$ is the coefficient for primary distortion.

$$W = W_{040}(\vec{\rho}\cdot\vec{\rho})^2 + W_{131}(\vec{H}\cdot\vec{\rho})(\vec{\rho}\cdot\vec{\rho}) + W_{222}(\vec{H}\cdot\vec{\rho})^2 + W_{220}(\vec{H}\cdot\vec{H})(\vec{\rho}\cdot\vec{\rho}) + W_{311}(\vec{H}\cdot\vec{H})(\vec{H}\cdot\vec{\rho}) \quad (2)$$

Buchroeder [30] has proven that the aberration field at the image plane of a non-symmetric optical system is still the sum of individual surface contributions, which are centered along different lines. No new aberrations are created in a system with a circular pupil or a quasi-circular pupil, but the nodal behavior will be exhibited. A vector $\vec{\sigma}_j$ is introduced, which points to the projection of the center of the aberration field of surface j onto the image plane. This vector is nothing but the decentration of the center of the aberration field.

The effective aberration field height of surface j is then written as

$$\vec{H}_{Aj} = \vec{H} - \vec{\sigma}_j \quad (3)$$

Replacing $\vec{H}$ with $\vec{H}_{Aj}$ into Eq. (1), the complete fourth-order wave aberration expansion for a non-symmetric system is given as follows as the summarization over surfaces.

$$\begin{aligned} W = {} & \Delta W_{20}(\vec{\rho}\cdot\vec{\rho}) + \Delta W_{11}(\vec{H}\cdot\vec{\rho}) + \sum_j W_{040_j}(\vec{\rho}\cdot\vec{\rho})^2 \\ & + \sum_j W_{131_j}[(\vec{H}-\vec{\sigma}_j)\cdot\vec{\rho}](\vec{\rho}\cdot\vec{\rho}) + \sum_j W_{222\,j}[(\vec{H}-\vec{\sigma}_j)\cdot\vec{\rho})^2 \\ & + \sum_j W_{220\,j}[(\vec{H}-\vec{\sigma}_j)\cdot(\vec{H}-\vec{\sigma}_j)](\vec{\rho}\cdot\vec{\rho}) \\ & + \sum_j W_{311\,j}[(\vec{H}-\vec{\sigma}_j)\cdot(\vec{H}-\vec{\sigma}_j)][(\vec{H}-\vec{\sigma}_j)\cdot\vec{\rho}] \end{aligned} \quad (4)$$

The defocus term $\Delta W_{20}(\vec{\rho}\cdot\vec{\rho})$ and tilt term $\Delta W_{11}(\vec{H}\cdot\vec{\rho})$ are included in the formula for completeness. Starting from the new wave aberration expansion, positions of nodal

points (points of zero aberration) of each primary aberration can be determined for specific systems. For example, the spherical aberration has no nodal points, while the primary coma has one node; the primary astigmatism has two nodes. Derivations of sixth-order wave aberrations are much more cumbersome in form provided in several publications [9-11].

Utilizing the nodal aberration theory, aberration field behavior that emerges with freeform surfaces can be analyzed, which is crucial for understanding the impacts of introducing freeform surfaces in an optical system. The aberration contributions resulted from some primary and higher order aberration terms of the freeform surface are investigated by Fuerschbach [14]. The main conclusion is, the freeform surface gives field constant net aberration contribution when placed at the pupil, and gives field dependent net aberration contribution when located away from the pupil. The emerging aberrations from some aberration terms are summarized in Table 1.

**Table 1.** Aberrations generated by freeform term overlay on an optical surface based on nodal aberration theory.

| **Freeform term (Zernike fringe coeff.)** | **Generated aberrations** | |
|---|---|---|
| | **At the stop** | **Away from the stop** |
| Zernike astigmatism (Z5, Z6) | Field constant astigmatism | Field constant astigmatism |
| Zernike coma (Z7, Z8) | Field constant coma | Field constant coma,<br>field linear astigmatism,<br>field linear medial field curvature |
| Zernike trefoil (Z10, Z11) | Field constant,<br>elliptical coma (trefoil) | Field linear astigmatism,<br>field constant elliptical coma (trefoil) |
| Zernike oblique spherical aberration (Z12, Z13) | Field constant, oblique spherical aberration | Field constant oblique spherical aberration,<br>field linear elliptical coma,<br>field linear coma,<br>field quadratic astigmatism,<br>field quadratic medial field curvature, |
| Zernike sixth order aperture coma (Z14, Z15) | Field constant sixth order aperture coma | Field constant sixth order aperture coma,<br>field linear medial oblique spherical aberration,<br>field linear oblique spherical aberration,<br>field quadratic trefoil,<br>hybrid field quadratic coma,<br>hybrid field cubic astigmatism,<br>field cubic medial field curvature |

| | | |
|---|---|---|
| Zernike tetrafoil (Z17, Z18) | Field constant tetrafoil | Field constant tetrafoil,<br>field linear trefoil,<br>field quadratic astigmatism |

As can be seen from this table, aberrations emerged from the existing freeform terms are rather complicated. Moreover, the freeform topology generally contains many terms, including many high order terms to achieve the best correction results. The interplay between different terms, and between different surfaces are rather cumbersome if the analytical solutions are used in multi-surface system design.

Oleszko [26] developed the tool based on Sasian's aberration theory [31] regarding induced and intrinsic aberrations in 2017, which is able to give the surface-resolved contribution of higher order aberrations numerically. Auxiliary entrance and exit pupils are inserted manually for each surface to track the wavefront deformation after each surface. This tool helps to understand the aberration compensation mechanism of a system, but has the limitation of only allowing for small fields due to the vignetting effects emerging at the edge of surfaces, when system pupils are decentered to cater off-axis fields.
Due to the redistribution of aberrations over multiple surfaces with the implementation of freeforms, a numerical investigation is indispensable.

### 2.2.2 ALDIS THEOREM

The Aldis theorem was introduced by Aldis, and first published by Cox in his book [1]. Unlike primary aberration expression which only gives first few orders of the whole transverse aberration, Aldis theorem gives surface contributions to the complete aberration of one ray including all orders. This property makes Aldis theorem quite handy in optical systems with high numerical apertures or large object fields, where higher order aberrations are unneglectable.

The original Aldis theorem was derived for rotationally symmetric spherical systems, where the image plane coincides with the Gaussian image plane. Brewer extended Aldis' equations to general, rotationally symmetric asphere [5, 6].

$$\mathrm{U_s} = \frac{\mu_n X_s}{b_n n_n}\left(\frac{b_s n_s}{\mu_s} - \frac{b_{s-1} n_{s-1}}{\mu_{s-1}}\right) - \frac{\mu_n Z_s}{b_n n_n}\left(\frac{b_s l_s}{\mu_s} - \frac{b_{s-1} l_{s-1}}{\mu_{s-1}}\right) - \frac{\mu_n \beta_s}{b_n n_n}\left(l_s - l_{s-1}\right) \qquad \textbf{(5)}$$

$$\begin{aligned} V_\mathrm{s} = {} & \frac{\mu_n Y_s}{b_n n_n}\left(\frac{b_s n_s}{\mu_s} - \frac{b_{s-1} n_{s-1}}{\mu_{s-1}}\right) - \frac{\mu_n Z_s}{b_n n_n}\left(\frac{b_s m_s}{\mu_s} - \frac{b_{s-1} m_{s-1}}{\mu_{s-1}}\right) \\ & - \frac{\mu_n \beta_s}{b_n n_n}\left(m_s - m_{s-1}\right) - \frac{\mu_n Y_p}{n_n}\left(\frac{n_s}{\mu_s} - \frac{n_{s-1}}{\mu_{s-1}}\right) \end{aligned} \qquad \textbf{(6)}$$

Where $(l_s, m_s, n_s)$ are optical direction cosines of ray after refraction at surface s; $(X_s, Y_s, Z_s)$ are Cartesian coordinates of the ray at surface s; $\beta_s$ is the height of paraxial marginal ray at the tangential plane (y-z plane) at surface s; $b_s$ is the paraxial marginal ray angle at the tangential plane at surface s times $\mu_s$; $\mu_s$ is the refractive index following surface s; the n subscript denotes image space, and $Y_p$ is the Gaussian image height at the tangential plane. $U_s$ and $V_s$ are contributions of individual surface s, so the overall aberrations can be characterized as

$$X_n = \sum_{s=1}^{n} U_s \qquad \textbf{(7)}$$

$$\Delta Y = Y_n - Y_p = \sum_{s=1}^{n} V_s \qquad \textbf{(8)}$$

With the extension from spheres to aspheres achieved by Brewer, the Aldis theorem is still limited to rotationally symmetric systems with only y fields, whose image planes are located at Gaussian image planes.

There have been very few visits and nearly no new development of the Aldis theorem since Brewer [32-34]. In the work by Chen and Herkommer [34], a differential ray pair is propagated numerically through the system using the transformation matrix to quantify individual surface aberration contributions, which gives similar results as the original Aldis theorem. However, the direct

generalization of the Aldis theorem to freeform surfaces, to x fields, to an arbitrary image plane, and to 3D system geometry is still needed.

## 2.3 FREEFORM SURFACES

Differing from the rotationally symmetric surfaces, the decoupling in the sagittal and tangential directions of freeform surfaces provides more possibilities for aberration correction than traditional surfaces.

### 2.3.1 DEFINITION

In general, the term freeform refers to surfaces with arbitrary topology. In contrast to illumination design, where freeform surfaces are more often described numerically due to their 'free' shapes, in imaging optical design, we prefer a more specific classification by using the term freeform for surfaces without rotational symmetry. This classification excludes a rotationally symmetric asphere placed in an off-axis position, which is toleranced and manufactured like a freeform. The principle of using the sub-aperture domain of an asphere is illustrated in Figure 1(b). In this case, although the aspheric surface representation is used for the surface, the sub-aperture domain still has no rotational symmetry [67]. Though being manufactured just like a freeform, the sub-aperture domain use of the asphere is less tricky in alignment compared with the freeform surfaces defined above [1]. These types of surfaces are referred to as aspheres in the context of this work.

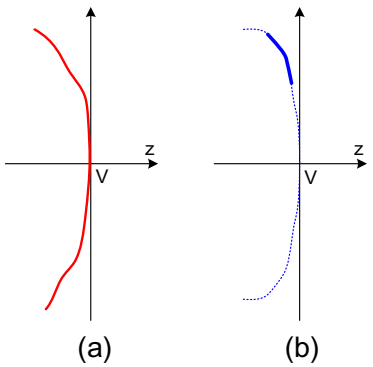



**Figure 1.** Illustration of off-axis sub-aperture domain. (a) is a non-rotationally symmetric freeform surface, (b)illustrates the sub-aperture domain use of a rotationally symmetric asphere.

### 2.3.2 FREEFORM SURFACE REPRESENTATIONS

Categorized by different types of orthogonality, different coordinate systems, and different weight and boundary functions, many globally defined freeform surface representations are available. They are all predefined sets of mathematical polynomials that can be utilized to describe surface shapes.

The general description of a freeform can be decomposed into a basic shape and a freeform deformation, which is formally expressed as [24]

$$z(x,y)=Z_{basic}+\frac{A(\bar{x},\bar{y})}{P(x,y)}\sum \mathrm{F}\,(\bar{x},\bar{y})\,, \qquad \textbf{(9)}$$

where $A(\bar{x},\bar{y})$ is the boundary function, $P(x,y)$ is the projection factor, and $\mathrm{F}\,(\bar{x},\bar{y})$ is the freeform term expressed using normalized Cartesian coordinates $\bar{x},\bar{y}$. $z(x,y)$ is the complete surface sag along z-axis.

The basic shape of the freeform contains the second order contribution of the surface, can have the shape of a sphere, a conic or a biconic. It is the dominating part in the neighborhood of the z-axis. In the case of a biconic basic shape, the astigmatic correction is introduced without using freeform deformation, which can provide a good pre-correction for systems with large astigmatism.

The boundary function defines the values of the deformation terms on the boundary of the surface, which can be used to control the surface center and edge values. The projection factor $P(x,y)$ provides the possibility of measuring the surface sag deviation along the local surface normal of the basic shape surface, which in principle links the deformation with the basic shape.

Many freeform surface representations are already available in commercial optical design software like Zemax (version 13, Kirkland, Washington, USA), which is selected as the software for this thesis. The non-orthogonal monomials representation (x-y polynomial), the spatial orthogonal Zernike representations (standard convention and fringe convention) and the gradient orthogonal Q-

polynomials are widely used freeform descriptions in the application. The first two types are supported directly in Zemax, while the last description can be implemented via user-defined DLL files. Of these representations, the non-orthogonal monomials show significant disadvantages in convergence compared with other orthogonal descriptions [25], is therefore not recommended. The variation between slope orthogonal and gradient orthogonal descriptions shows no significant difference in convergence and final system performance according to Brömel et al. [25].

The Zernike fringe representation is defined in polar coordinates and has a constant weighting function [35].

$$z(\bar{r},\varphi)=\sum_{m,n} a_n^m Z_n^m(\bar{r},\varphi) \quad \text{with m,n=0,1...N} \tag{10}$$

$a_n^m$ is the coefficient of individual Zernike term $Z_n^m(\bar{r},\varphi)$. $Z_n^m(\bar{r},\varphi)$ is the product of the radial part and the azimuthal part,

$$Z_n^m(\bar{r},\varphi)=R_n^m(\bar{r})\,\Phi^m(\varphi) \tag{11}$$

which can be written as

$$\begin{aligned} R_n^m(\bar{r}) &= \sum_{k=0}^{\frac{n-|m|}{2}} (-1)^k \binom{n-m}{k}\binom{n-2k}{\frac{n-m}{2}-k}\bar{r}^{n-2k} \\ &= \sum_{k=0}^{\frac{n-|m|}{2}} (-1)^k \frac{(n-k)!}{k!\left(\frac{n-|m|}{2}-k\right)\left(\frac{n+|m|}{2}-k\right)}\bar{r}^{n-2k} \end{aligned} \tag{12}$$

and

$$\Phi^m(\varphi)=\begin{cases}\sin(m\varphi), & m<0\\ \cos(m\varphi), & m>0\\ 1, & m=0\end{cases} \tag{13}$$

### 2.3.3 FREEFORM MANUFACTURABILITY

The Q type asphere is first introduced by G.W. Forbes in 2007 [36] as an alternative representation assisting manufacturability assessment. This asphere is later generalized for the application of freeform optics [37], and being referred to as Q-polynomials. Sets of orthogonal polynomials are introduced to characterize surface sag deviation from the best-fit sphere. This freeform surface representation is defined on a polar coordinate system with slope orthogonality, which means the gradient of the surface departure is composed of only orthogonal terms. The Q-polynomial representation is shown in Eq. (14).

$$z(\rho,\theta)=\frac{c\rho^2}{1+\sqrt{1-c^2\rho^2}}+\frac{1}{\sqrt{1-c^2\rho^2}}\left\{u^2(1-u^2)\sum_{n=0}^{N}a_n^0Q_n^0(u^2)+\sum_{m=1}^{M}u^m\sum_{n=0}^{N}[a_n^m\cos(m\theta)+b_n^m\sin(m\theta)]Q_n^m(u^2)\right\} \quad (14)$$

In Eq. (14), c is the curvature of the "best-fit sphere", $\sqrt{1-c^2\rho^2}$ is the projection function, which projects the surface departure along the local surface normal. $a_n^m$, $b_n^m$ are coefficients of freeform surface terms, and $Q_n^m(v)$ is a polynomial of order n in $v$.

Because the manufacturability is often driven by the rate of change of the normal departure from the best-fit sphere, this slope orthogonal representation can be conveniently related to surfaces' testability. The freeform deformation part of Eq. (14) is shown in Eq. (**15**), whose mean square slope satisfies Eq. (**16**).

$$\delta(u,\theta)\coloneqq u^2(1-u^2)\sum_{n=0}^{N}a_n^0Q_n^0(u^2)+\sum_{m=1}^{M}u^m\sum_{n=0}^{N}[a_n^m\cos(m\theta)+b_n^m\sin(m\theta)]Q_n^m(u^2) \quad (15)$$

$$\langle|\nabla\delta(\mu,\theta)|^2\rangle = \langle\left(\frac{\partial\delta}{\partial u}\right)^2 + \frac{1}{u^2}\left(\frac{\partial\delta}{\partial\theta}\right)^2\rangle = \sum_{m,n}[(a_n^m)^2 + (b_n^m)^2] \qquad \textbf{(16)}$$

Since the root mean square (RMS) fringe density of the resulting surface interferogram is related to the RMS of the departure slope, Eq. (**16**) is helpful for estimating freeforms' testability. Although it is the peak fringe density that drives testability, the RMS fringe density is more accessible and can be used as a measure.

According to [1, 38], the freeform manufacturing constraints given for a Moore Nanotechnology 450UPL machine are:

- Freeform tool stroke≤6mm
- Azimuthal gradients≤10°
- Azimuthal accelerations ≤ 5g

The restrictions above are given for the non-rotationally symmetric freeform part of the surface data, which is decomposed from the complete data set along with a rotational symmetric part. These requirements set limits for the freeform topology that ultimately determine freeforms' manufacturability.

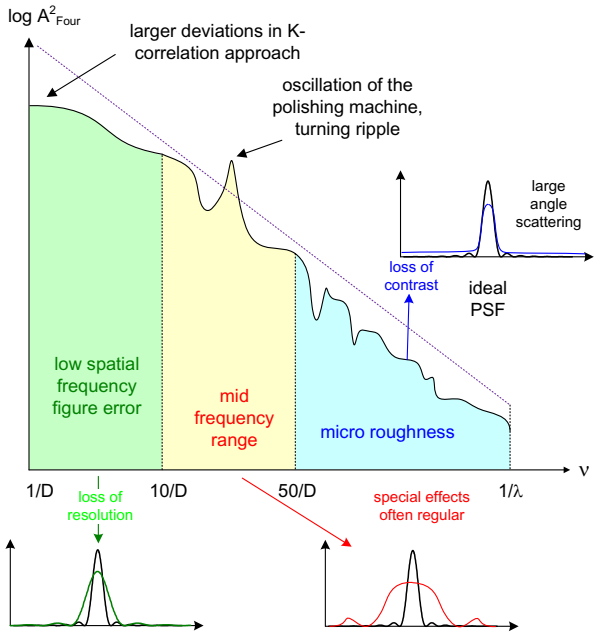



**Figure 2.** Illustration of the power spectral density of the freeform surface [3].

Freeform errors coming from the manufacturing process are categorized into three regimes: the low spatial frequency figure error responsible for the loss of resolution, the mid-frequency error resulted from the regular diamond-turning ripples, and the high spatial frequency surface microroughness leading to the loss of contrast [2].

### 2.3.4 FREEFORM ORDERING AND SYMMETRY

Ordering and symmetry information of the freeform surface representations Zernike fringe and Q-polynomial up to the $10^{th}$ order are given in Table 17 and Table 18 respectively in the Appendix. This information is necessary for deciding freeform variables based on the optical system symmetry and aberration constitution.

## 2.4 EXAMPLE OPTICAL SYSTEMS

Several example optical system types are introduced in this section, laying the foundation for the practical check of the theoretical results in the following work.

### 2.4.1 SCHEIMPFLUG CONDITION

When the object plane of the optical system is tilted by a certain angle $\theta$, a sharp image is only possible when the image plane has a tilt angle of $\theta'$. This is being referred to as the paraxial Scheimpflug condition [39-42]. As explained with a simple layout in Figure 3, the extension of the image plane and the object plane intercept at the principal plane of the system under the Scheimpflug condition.

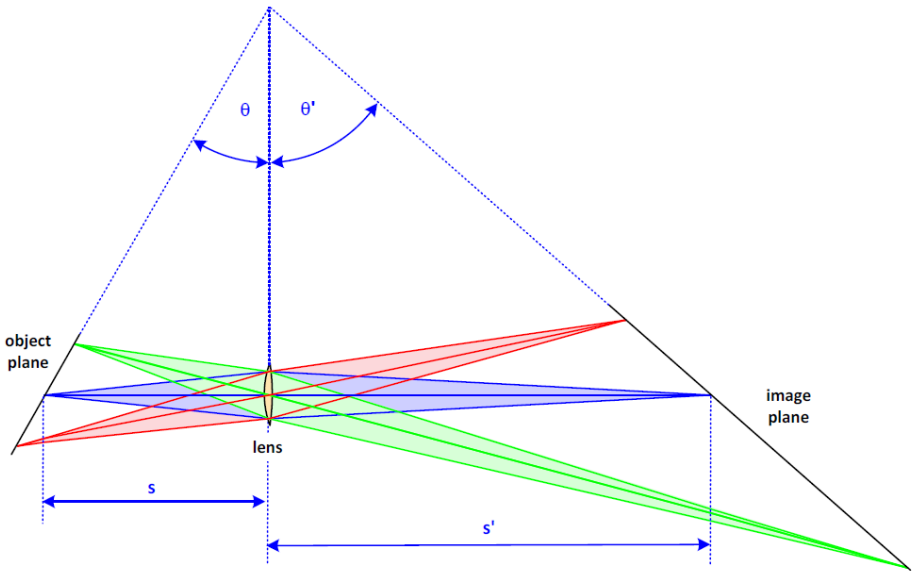



**Figure 3.** Indication of the scheimpflug condition in a simple optical system. [39]

The magnification of the object point on the optical axis can be written as:

$$m_0 = \frac{s'}{s} = \frac{\tan\theta'}{\tan\theta} \tag{17}$$

Due to the existence of the tilted object and image planes, the local magnification varies for different object heights; correspondingly, a large keystone distortion exists. The anamorphic magnifications in the sagittal and meridional sections caused by the Scheimpflug condition are:

$$m_x = m_0 \tag{18}$$

$$m_y = m_0^2 \cdot \frac{\sin\theta}{\sin\theta'} \tag{19}$$

The Scheimpflug system finds its wide application in fields like triangulation sensor [43], 3D measurement and so on.

### 2.4.2 ALVAREZ LENS SYSTEM

The Alvarez lens pair is composed of two closely placed glass plates, whose inner surfaces share the same cubic freeform shape. The classical Alvarez lens pair has the

following sag function [39, 44], where A is the amplitude of the individual freeform surface.

$$z(x,y)=A\left(xy^2+\frac{1}{3}x^3\right) \tag{20}$$

Here the Alvarez lens pair is represented using the monomials representation, which can be sorted by aberrational order similar to the Zernike fringe representation [35]. The classical Alvarez lens pair is therefore of 6th order. If both Alvarez plates are shifted by $\Delta x$ in opposite directions along the x-axis as illustrated in Figure 4, the resulting lens optical thickness is

$$\begin{aligned} z^{(x)}(x,y) &= z(x+\Delta x,y)-z(x-\Delta x,y) \\ &= 2A\Delta x\left[x^2+y^2\right]+2A\Delta x^3 \end{aligned} \tag{21}$$

The last term of the above equation is a constant, whereas the first term corresponds to defocus.

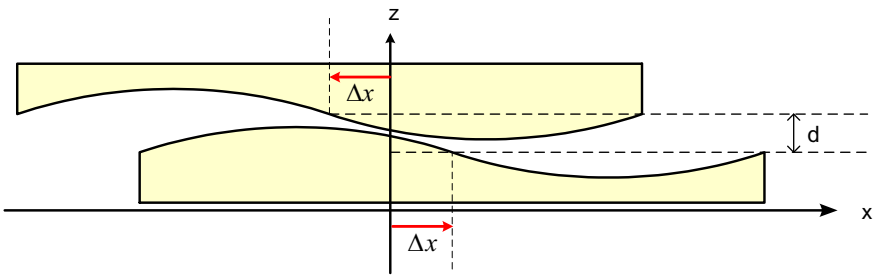



**Figure 4.** Defocus concept of the Alvarez lens pair.

If the shifts of both plates are performed along the y-axis, an astigmatism effect will be obtained as indicated in Eq. (22).

$$z^{(y)}(x,y)=-4Axy\Delta y \tag{22}$$

The defocusing property of the Alvarez lens can be used for zoom systems without the need to vary the total length of the system. In reality, due to the finite distance needed between two Alvarez plates to avoid the collision, the induced aberrations originated at the second plate dominate. The wavefront deviations coming from the

first plate are responsible for the induced aberrations, whose magnitudes are related to the freeform amplitude A. The larger the amplitude A, the larger the freeform sag, the smaller the displacement needed to reach the same dynamic focal range, and the larger the induced aberrations. Due to the mostly used thin element approach of the Alvarez lens pair, both freeforms can be placed either at the inner surfaces or at the outer surfaces. The difference between these two cases is the actual distance separating two freeforms. With freeforms located at the outer surfaces, the induced aberrations are of larger magnitude due to increased spacing d.

The induced aberrations can be calculated analytically, and only certain higher order terms are induced based on the low order surface representation. The overall wavefront aberration of the Alvarez lens pair is obtained through subtracting the optical path differences of two plates after shifting $\Delta x$ in x. The overall wavefront deformation considering the Alvarez lens spacing d can be written as follows [45].

$$
\begin{aligned}
&W(x,y) = dA^2(n-1)^2(-x^4-6x^2y^2-y^4)+d^2A^3(n-1)^3(-x^5-10x^3y^2-5xy^4) \\
&+d^3A^4(n-1)^4(-\frac{1}{3}x^6-5x^4y^2-5x^2y^4-\frac{1}{3}y^6)+ \\
&\Delta x[A(n-1)(2x^2+2y^2)+A^3(n-1)^3(-3x^4-18x^2y^2-3y^4) \\
&+dA^4(n-1)^3(-4x^7-44x^5y^2-60x^3y^4-20xy^6) \\
&+d^2A^3(n-1)^2(-2x^6-14x^4y^2-14x^2y^4-2y^6) \\
&+d^2A^5(n-1)^4(-2x^8-32x^6y^2-60x^4y^4-32x^2y^6-2y^8) \\
&+d^3A^4(n-1)^4(-2x^5-20x^3y^2-10xy^4)]+[...]
\end{aligned}
\tag{23}
$$

Only certain aberrations exist in the above equation, which means, only these terms are needed in the freeform representation for induced aberration correction. Of these terms, aberrations belonging to the 8th order are terms $x^4$, $x^2y^2$, $y^4$.

Maximum lens pair displacement is critical for the magnitudes of induced aberrations, which is limited by the mechanical method used. A reasonable displacement range is 200μm for a 2mm diameter lens [46].

### 2.4.3 SPECTROMETER SYSTEM

A spectrometer system is able to provide additionally objects' spectral information other than the spatial information. A typical spectrometer is composed of three parts, a collimator that collimates the input field, a dispersive element that separates spectral information and an imager that focuses light onto the detector. The spectrometer still works without a collimator [47], but the varying incidence angles on the disperser due to the absence of collimator leads to additional degradation of the spectral resolution and system efficiency [48].

Typical dispersive elements are prism and grating, which have the opposite dispersion properties. The prism gives crowded red fringes, while the grating gives crowded blue fringes. A proper combination of the prism and grating is called grism, and is able to provide linear dispersion [49-52]. Typical mirror systems as Offner [53], Schwarzschild [54], and three mirror anastigmatic (TMA) [55], are used widely for spectrometer designs [56-58]. The concentric configuration of the first two system types provides a simple layout with minimum residual aberrations as the start system. The TMA system presents the possibility of using the double pass layout, which utilizes the same system twice for the collimator and the imager. By applying the double pass configuration on the TMA system with perpendicular incidence on the grating, the complexity of the original double TMA system is reduced greatly with an improved resolution due to the quasi-symmetric layout. A hidden advantage of using a mirror system is the vanished color aberration, which is a superior property when it comes to the application of a wide spectral range. In particular, the implementation of freeform surfaces helps to fulfill the 3D correction requirement of the spectrometer system, as well as present more compact designs in the case of mirror systems.

The catadioptric Dyson system is also a common compact spectrometer design, whose quasi-symmetric layout is advantageous for many primary aberration corrections. In the work of this thesis, the focus is mainly on the design of the grating spectrometer systems due to their constant resolving power and compact configurations.

### 2.4.4 OTHER SYSTEMS

Other than the system types mentioned above, there are some more systems that can also benefit from the use of freeform surfaces. One example is the non-rotationally symmetric lithography system. When mirror elements are introduced in such a lithography system for correction purposes [65], the original rotational symmetry is broken, which makes the introduction of freeform surface beneficial.

Another example is the telecentric two-mirror scan system. The deflection of one mirror in both scanning directions is hard to realize in reality, therefore two scanning mirrors with a respective deflection in x and y are needed for an x-y scan system. The scanning system should also be telecentric on the image side. For the traditional rotationally symmetric scan system, the pupil locations for x-z and y-z planes coincide, meaning only one scanning mirror can be placed at the exact pupil location. With the other scanning mirror shifted from the pupil location, the telecentricity requirement is violated. Freeform is useful for such an application, with whose help the separation of both pupil positions becomes possible.

One more example is a reflective zoom system. Traditional zoom systems are typically constructed with moving refractive elements, the concept of an all-reflective zoom system can be realized with the help of freeforms [70, 71]. The occurring large off-axis aberrations due to decentered and tilted mirror components can be compensated better with the aberration decoupling property of freeform surfaces. Applications like zoom telescope systems become possible.

# 3 Extended Aldis theorem

As mentioned in section 2.2.2, the Aldis theorem presents surface aberration contributions including all orders for one real ray. This unique property of Aldis theorem makes it advantageous for freeform systems evaluation, given the generalization to more complicated systems can be realized. Starting from Brewer's interpretation of the Aldis theorem, the extension of his work to more generalized surface types and system types will be conducted.

## 3.1 THEOREM DERIVATION FOR A FINITE X FIELD

The original theorem only allows for y fields in the investigated system. The derivation of the surface aberration contribution equations including the x field starts from two ray tracing equations Eq. (**24**) and (**25**). $(X_s, Y_s, Z_s)$ are real ray intersection coordinates of the target ray at surface s, $(l_s, m_s, n_s)$ are the optical direction cosines of this ray after refraction at surface s. $\mu_s$ represents refractive index after surface s, $\beta x_s$ is the paraxial marginal ray height at the sagittal plane of surface s, $bx_s$ is the paraxial marginal ray angle after refraction at the sagittal plane of surface s times $\mu_s$, and $t_{s-1}$ represents the axial distance between surface s-1 and surface s. These symbols are illustrated in Figure 5.

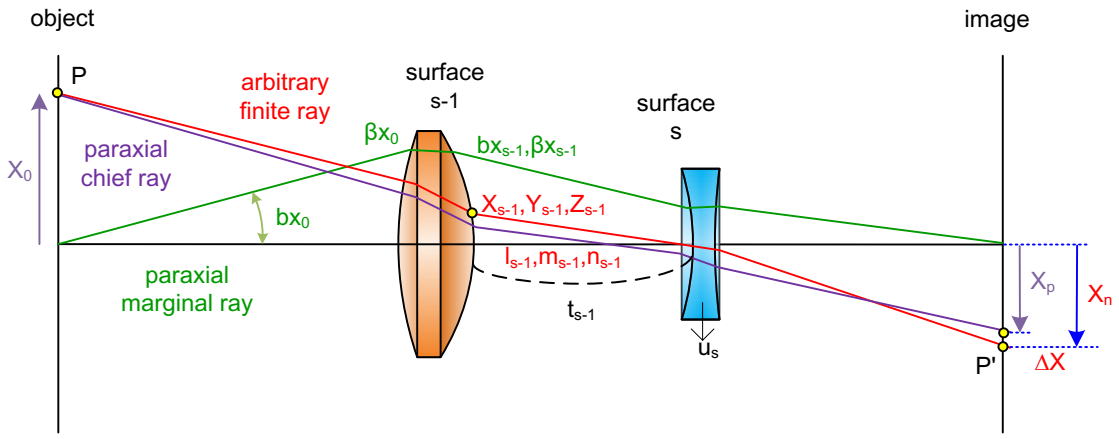



**Figure 5.** Quantities used for Aldis theorem derivation.

The two ray tracing equations include no information of the surface shape, and therefore can be applied for general surfaces including freeforms.

$$\frac{X_s - X_{s-1}}{l_{s-1}} = \frac{Y_s - Y_{s-1}}{m_{s-1}} = \frac{Z_s - Z_{s-1} + t_{s-1}}{n_{s-1}} \tag{24}$$

$$\beta x_s = \beta x_{s-1} + bx_{s-1} \cdot \frac{t_{s-1}}{\mu_{s-1}} \tag{25}$$

Multiply the first and the third term of Eq. (**24**) by $\frac{bx_{s-1}}{\mu_{s-1}}$, we get

$$\frac{(X_s - X_{s-1}) \cdot n_{s-1} \cdot bx_{s-1}}{\mu_{s-1}} = \frac{(Z_s - Z_{s-1}) \cdot l_{s-1} \cdot bx_{s-1}}{\mu_{s-1}} + \frac{t_{s-1} \cdot bx_{s-1}}{\mu_{s-1}} \cdot l_{s-1}. \tag{26}$$

The term $\frac{t_{s-1} \cdot bx_{s-1}}{\mu_{s-1}}$ can be substituted from Eq. (**25**).

$$\frac{(X_s - X_{s-1}) \cdot n_{s-1} \cdot bx_{s-1}}{\mu_{s-1}} = \frac{(Z_s - Z_{s-1}) \cdot l_{s-1} \cdot bx_{s-1}}{\mu_{s-1}} + (\beta x_s - \beta x_{s-1}) \cdot l_{s-1} \tag{27}$$

$$\frac{X_s \cdot n_{s-1} \cdot bx_{s-1}}{\mu_{s-1}} - \frac{X_{s-1} \cdot n_{s-1} \cdot bx_{s-1}}{\mu_{s-1}} = \frac{Z_s \cdot l_{s-1} \cdot bx_{s-1}}{\mu_{s-1}} - \frac{Z_{s-1} \cdot l_{s-1} \cdot bx_{s-1}}{\mu_{s-1}} + (\beta x_s - \beta x_{s-1}) \cdot l_{s-1} \tag{28}$$

Adding term $\beta x_s \cdot (l_s - l_{s-1})$ on both sides, Eq. (**28**) can be written as

$$-\frac{X_{s-1}\cdot n_{s-1}\cdot bx_{s-1}}{\mu_{s-1}}+\frac{Z_{s-1}\cdot l_{s-1}\cdot bx_{s-1}}{\mu_{s-1}}-\beta x_s\cdot l_s+\beta x_{s-1}\cdot l_{s-1}=$$
$$-\frac{X_s\cdot n_{s-1}\cdot bx_{s-1}}{\mu_{s-1}}+\frac{Z_s\cdot l_{s-1}\cdot bx_{s-1}}{\mu_{s-1}}-\beta x_s\left(l_s-l_{s-1}\right). \quad \textbf{(29)}$$

Adding $\frac{X_s\cdot n_s\cdot bx_s}{\mu_s}-\frac{Z_s\cdot l_s\cdot bx_s}{\mu_s}$ to both sides of the Eq. (**29**) gives

$$\left(\frac{X_s\cdot n_s\cdot bx_s}{\mu_s}-\frac{X_{s-1}\cdot n_{s-1}\cdot bx_{s-1}}{\mu_{s-1}}\right)-\left(\frac{Z_s\cdot l_s\cdot bx_s}{\mu_s}-\frac{Z_{s-1}\cdot l_{s-1}\cdot bx_{s-1}}{\mu_{s-1}}\right)-\left(\beta x_s\cdot l_s-\beta x_{s-1}\cdot l_{s-1}\right)=$$
$$X_s\left(\frac{n_s\cdot bx_s}{\mu_s}-\frac{n_{s-1}\cdot bx_{s-1}}{\mu_{s-1}}\right)-Z_s\left(\frac{l_s\cdot bx_s}{\mu_s}-\frac{l_{s-1}\cdot bx_{s-1}}{\mu_{s-1}}\right)-\beta x_s\left(l_s-l_{s-1}\right). \quad \textbf{(30)}$$

Summing both sides of Eq. (**30**) for $s=1$ to $s=n$ gives

$$\left(\frac{X_n\cdot n_n\cdot bx_n}{\mu_n}-\frac{X_0\cdot n_0\cdot bx_0}{\mu_0}\right)-\left(\frac{Z_n\cdot l_n\cdot bx_n}{\mu_n}-\frac{Z_0\cdot l_0\cdot bx_0}{\mu_0}\right)-\left(\beta x_n\cdot l_n-\beta x_0\cdot l_0\right)=$$
$$\sum_{s=1}^{n}\left[X_s\left(\frac{n_s\cdot bx_s}{\mu_s}-\frac{n_{s-1}\cdot bx_{s-1}}{\mu_{s-1}}\right)-Z_s\left(\frac{l_s\cdot bx_s}{\mu_s}-\frac{l_{s-1}\cdot bx_{s-1}}{\mu_{s-1}}\right)-\beta x_s\left(l_s-l_{s-1}\right)\right]. \quad \textbf{(31)}$$

Assuming planar object and image without surface tilts, then $Z_n=Z_0=0$, and the third and fourth terms of the left hand side of Eq. (**31**) becomes zero. $\beta x_0=0$, $\beta x_n$ also equals to zero for the Gaussian image plane, the fifth and sixth terms are also eliminated. The left hand side of Eq. (**31**) can then be written as

$$\frac{X_n\cdot n_n\cdot bx_n}{\mu_n}-\frac{X_0\cdot n_0\cdot bx_0}{\mu_0}. \quad \textbf{(32)}$$

The Lagrange invariant of the optical system gives $X_p\cdot bx_n=X_0\cdot bx_0$, where $X_p$ is the Gaussian image height of the chief ray in x. The left hand side of Eq. (**31**) is further modified to

$$\frac{X_n \cdot n_n \cdot bx_n}{\mu_n} - \frac{X_p \cdot n_n \cdot bx_n}{\mu_n} + \frac{X_p \cdot n_n \cdot bx_n}{\mu_n} - \frac{X_p \cdot n_0 \cdot bx_n}{\mu_0} \tag{33}$$

$$= \frac{n_n \cdot bx_n}{\mu_n}\left(X_n - X_p\right) + X_p \cdot bx_n \sum_{s=1}^{n}\left(\frac{n_s}{\mu_s} - \frac{n_{s-1}}{\mu_{s-1}}\right). \tag{34}$$

Taking the above equation back into Eq. (**31**), we obtain

$$\begin{aligned}\frac{bx_n \cdot n_n}{\mu_n}\left(X_n - X_p\right) = \sum_{s=1}^{n}\Bigg[ & X_s\left(\frac{bx_s \cdot n_s}{\mu_s} - \frac{bx_{s-1} \cdot n_{s-1}}{\mu_{s-1}}\right) - Z_s\left(\frac{bx_s \cdot l_s}{\mu_s} - \frac{bx_{s-1} \cdot l_{s-1}}{\mu_{s-1}}\right) - \beta x_s\left(l_s - l_{s-1}\right)\Bigg] \\ & - bx_n \cdot X_p \sum_{s=1}^{n}\left(\frac{n_s}{\mu_s} - \frac{n_{s-1}}{\mu_{s-1}}\right).\end{aligned} \tag{35}$$

The transverse aberration in y direction remains the same as Brewer's equation, where $Y_p$ is the Gaussian image height of the chief ray in y. $\left(X_n - X_p\right)$ equals to the transverse aberration in x, and $\left(Y_n - Y_p\right)$ equals to the transverse aberration in y.

$$\begin{aligned}\frac{b_n \cdot n_n}{\mu_n}\left(Y_n - Y_p\right) = \sum_{s=1}^{n}\Bigg[ & Y_s\left(\frac{b_s \cdot n_s}{\mu_s} - \frac{b_{s-1} \cdot n_{s-1}}{\mu_{s-1}}\right) - Z_s\left(\frac{b_s \cdot m_s}{\mu_s} - \frac{b_{s-1} \cdot m_{s-1}}{\mu_{s-1}}\right) - \beta_s\left(m_s - m_{s-1}\right)\Bigg] \\ & - b_n \cdot Y_p \sum_{s=1}^{n}\left(\frac{n_s}{\mu_s} - \frac{n_{s-1}}{\mu_{s-1}}\right)\end{aligned} \tag{36}$$

The transverse aberrations in y and x are written as the summation of individual surface contributions for the target ray.

$$\Delta Y = Y_n - Y_p = \sum_{s=1}^{n} V_s, \Delta X = X_n - X_p = \sum_{s=1}^{n} U_s \tag{37}$$

## 3.2 EXTENSION TO ARBITRARY IMAGE PLANES

The original Aldis theorem is derived for an optical system with the image formed at the Gaussian image plane. However, the real image plane is more likely shifted in reality to have a minimum spot radius. Extra defocus of the system is introduced in

this case by moving the image plane away from the Gaussian image plane. In the case of an arbitrary image plane, $\beta x_0 = 0$ is still valid for the object, but the paraxial marginal height $\beta x_n$ is no longer zero at the image plane, the term $\beta x_n \cdot m_n$ can be treated as one of the two components of system defocus.

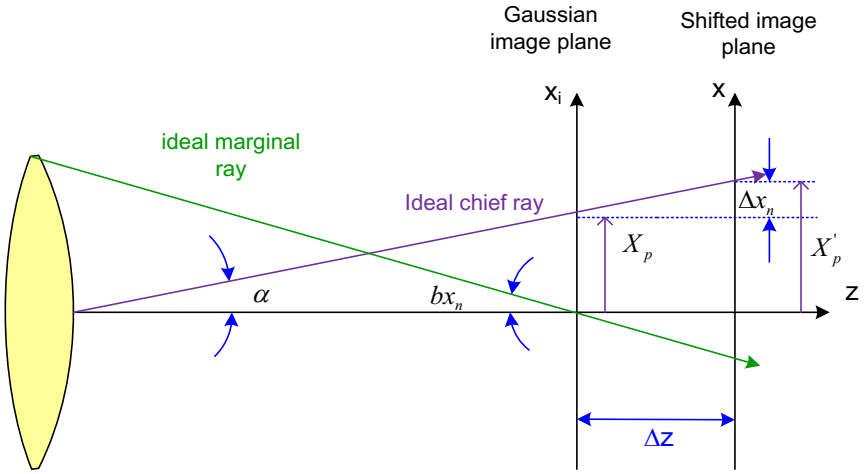



**Figure 6.** Extension of Aldis theorem to an arbitrary image plane. The reference ray height given by the chief ray is different due to the displacement of the image plane.

The second component of the defocus originates from Eq. (**32**) due to the change of intersection height of the paraxial chief ray at the image plane. As illustrated in Figure 6, where the Gaussian image height is $X_p$, and the paraxial image height at the shifted image plane is $X_p{}'$. The Lagrange invariant is $X_p \cdot bx_n = X_0 \cdot bx_0$ for the Gaussian image plane. The new paraxial chief ray height at the shifted image plane is then $X_p{}' = X_p + \Delta x_n$. Taking this into Eq. (**32**) yields

$$\frac{X_n \cdot n_n \cdot bx_n}{\mu_n} - \frac{n_0 \cdot bx_n}{\mu_0}\left(X_p{}' - \Delta x_n\right). \qquad \textbf{(38)}$$

$$\frac{X_n \cdot n_n \cdot bx_n}{\mu_n} - \frac{X_p{}' \cdot n_n \cdot bx_n}{\mu_n} + \frac{X_p{}' \cdot n_n \cdot bx_n}{\mu_n} - \frac{X_p{}' \cdot n_0 \cdot bx_n}{\mu_0} + \frac{n_0 \cdot bx_n \cdot \Delta x_n}{\mu_0} \qquad \textbf{(39)}$$

$$= \frac{n_n \cdot bx_n}{\mu_n}\left(X_n - X_p^{'}\right) + bx_n \cdot X_p^{'} \sum_{s=1}^{n}\left(\frac{n_s}{\mu_s} - \frac{n_{s-1}}{\mu_{s-1}}\right) + \frac{n_0 \cdot bx_n \cdot \Delta x_n}{\mu_0} \qquad \textbf{(40)}$$

The final aberration in x direction is written as Eq. (**41**), where the terms inside the last parentheses represent defocus due to image plane displacement.

$$\begin{aligned}\frac{bx_n \cdot n_n}{\mu_n}\left(X_n - X_p^{'}\right) = \sum_{s=1}^{n}\Bigg[ & X_s\left(\frac{bx_s \cdot n_s}{\mu_s} - \frac{bx_{s-1} \cdot n_{s-1}}{\mu_{s-1}}\right) - Z_s\left(\frac{bx_s \cdot l_s}{\mu_s} - \frac{bx_{s-1} \cdot l_{s-1}}{\mu_{s-1}}\right) - \beta x_s\left(l_s - l_{s-1}\right)\Bigg] \\ & - bx_n \cdot X_p^{'} \sum_{s=1}^{n}\left(\frac{n_s}{\mu_s} - \frac{n_{s-1}}{\mu_{s-1}}\right) - \left(\frac{n_0 \cdot bx_n \cdot \Delta x_n}{\mu_0} - \beta x_n \cdot l_n\right)\end{aligned} \qquad \textbf{(41)}$$

The surface contributions also experience minor changes due to the change of the reference ray height $X_p^{'}$, which in most cases is still comparable to the Gaussian image height.

The x-z plane case is used above as example, while the derivation for the y-z plane follows the same philosophy. Similarly, the extra defocus term for the y direction is $-\left(\frac{n_0 \cdot b_n \cdot \Delta y_n}{\mu_0} - \beta_n \cdot m_n\right)$. To be noted, the Gaussian image plane varies for the x field and for the y field for systems with biconic surfaces due to different paraxial radii of curvature. Therefore $\Delta y_n$ and $\Delta x_n$ need to be calculated with respect to different image plane shift $\Delta z$.

## 3.3 MULTIPLE RAY TRACING

The other limitation of the Aldis theorem is, only transverse aberrations of one real ray are given by the equations, while the full pupil distribution of that specific field point can only be determined by tracing multiples rays with proper pupil sampling.

To get the complete picture of the individual surface contribution to one field point, the so-called peak-valley (PV) method is utilized, in which the maximum and minimum transverse aberrations of all traced rays are used to represent the extreme cases. When considering the calculated transverse aberrations on the spot diagram,

each point corresponds to one traced real ray, whose deviations in x and y with respect to the reference represent transverse aberrations. Though the field resolution is determined mostly by the RMS radius of the collective ray intersection points on the image plane, the PV value is also an important parameter for field resolution evaluation.

Due to the summation and subtraction nature of the PV calculation, the individual surface contribution for one field point can be obtained after multiple ray tracing. With a proper sampling of the pupil plane, this contribution can serve as a full order alternative for the Seidel diagram. The PV surface contribution for one field can be written as Eq. (**42**), where i represents traced real ray number with different pupil coordinates.

$$\begin{aligned} PV &= \Delta Y_{\max} - \Delta Y_{\min} = \sum_{s=1}^{n} V_{i_{\max},s} - \sum_{s=1}^{n} V_{i_{\min},s} \\ &= \sum_{s=1}^{n} \left( V_{i_{\max},s} - V_{i_{\min},s} \right) \end{aligned} \tag{42}$$

## 3.4 PARABASAL RAY TRACING

In the previous sections, the problem of extending the Aldis theorem on 2D systems with freeform surfaces has been solved. More general cases with 3D system geometry remain to be considered, where the common optical axis no longer exists.

In parabasal models [59-61], the optical system is evaluated around a base ray, which replaces the functionality of an optical axis. The Aldis theorem is extended from the paraxial region to the parabasal region for systems with 3D geometries.

Paraxial ray tracing is no longer applicable for systems with 3D geometries, instead, a new reference ray is defined connecting the center of the object and the center of the active stop. This ray is being referred to as the base ray, and rays located around the base ray are named as parabasal rays. Aberration and image assessment calculations based on the paraxial ray tracing can then be extended using parabasal ray tracing.

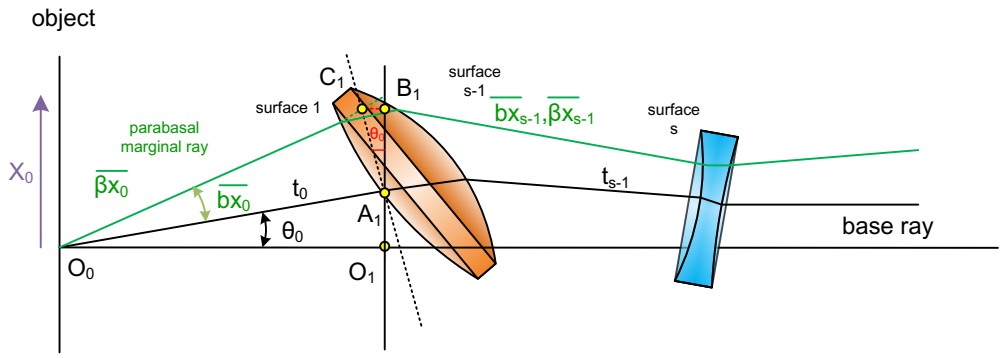



**Figure 7.** Illustration of the parabasal model for the Aldis theorem.

The symbols used for parabasal ray tracing are explained in Figure 7. Both lenses are tilted and decentered in the x-z plane, the system remains to be y symmetric in this case, but more general cases can be extended easily. The base ray no longer coincides with the z-axis, but form a finite angle $\theta_s$with it after surface s. Similar to paraxial ray tracing, the parabasal marginal rays and chief rays are evaluated with respect to the base ray and represented with bars above. The parabasal marginal ray height at surface s relative to the base ray is $A_sC_s$ for a small angle deviation $\overline{bx_{s-1}}$. The overall x coordinate height in the global coordinate system is $O_sB_s = O_sA_s + A_sB_s$, where $A_sB_s = A_sC_s \times \cos(\theta_{s-1})$. $A_sC_s$ can be scaled up from a ray very close to the base ray with small pupil coordinate difference. Similarly, the parabasal ray tracing for chief rays is also realized with respect to the real ray tracing coordinates of the base ray. Due to the decentration and tilt of the system elements, the base ray forms a finite angle with the z-axis, which leads to variations between upper marginal ray and lower marginal ray angles. Both rays are traced, whose average is used as a better approximation of the parabasal marginal ray angle in the case of large $\theta_{s-1}$.

The starting equations Eq.(**24**) and (**25**) for Aldis theorem derivation are written as follows for Figure 7 for the real ray tracing and the parabasal ray tracing.

$$\frac{X_s - X_{s-1}}{l_{s-1}} = \frac{Y_s - Y_{s-1}}{m_{s-1}} = \frac{Z_s - Z_{s-1} + t_{s-1} \cdot \cos\theta_{s-1}}{n_{s-1}} \qquad \textbf{(43)}$$

$$\overline{\beta x_s} = \overline{\beta x_{s-1}} + \overline{bx_{s-1}} \cdot \frac{t_{s-1} \cdot \cos\theta_{s-1}}{\mu_{s-1}} \qquad \textbf{(44)}$$

Similar to the derivation in section 3.1, both equations are connected through the new surface distance $t_{s-1}$. Quantities represented with respect to the base ray are indicated with a bar. The extra $\cos\theta_{s-1}$ term cancels out following the exact same procedure as in section 3.1, the same intermediate expression is obtained similar to Eq. (**31**) for the parabasal ray tracing.

$$\left(\frac{X_n\cdot n_n\cdot\overline{bx_n}}{\mu_n}-\frac{X_0\cdot n_0\cdot\overline{bx_0}}{\mu_0}\right)-\left(\frac{Z_n\cdot l_n\cdot\overline{bx_n}}{\mu_n}-\frac{Z_0\cdot l_0\cdot\overline{bx_0}}{\mu_0}\right)-\left(\overline{\beta x_n}\cdot l_n-\overline{\beta x_0}\cdot l_0\right)= \tag{45}$$
$$\sum_{s=1}^{n}\left[X_s\left(\frac{n_s\cdot\overline{bx_s}}{\mu_s}-\frac{n_{s-1}\cdot\overline{bx_{s-1}}}{\mu_{s-1}}\right)-Z_s\left(\frac{l_s\cdot\overline{bx_s}}{\mu_s}-\frac{l_{s-1}\cdot\overline{bx_{s-1}}}{\mu_{s-1}}\right)-\overline{\beta x_s}\left(l_s-l_{s-1}\right)\right]$$

The left hand side of Eq. (**45**) is written for the parabasal model as

$$\frac{X_n\cdot n_n\cdot\overline{bx_n}}{\mu_n}-\frac{X_0\cdot n_0\cdot\overline{bx_0}}{\mu_0}\quad, \tag{46}$$

assuming perpendicular object and image planes to the z axis. For this parabasal model, the real rays are represented in the global coordinate system, while the parabasal ray height and angle information are represented with respect to the base ray. If we use $X_s^{\,b}$ to represent the x intersection height of the base ray at surface s, the Lagrange invariant of the system gives $(X_p-X_n^b)\cdot\overline{bx_n}=(X_0-X_0^b)\cdot\overline{bx_0}$, with $X_p$ being the Gaussian image height of the chief ray in the global coordinate system as explained in Figure 8.

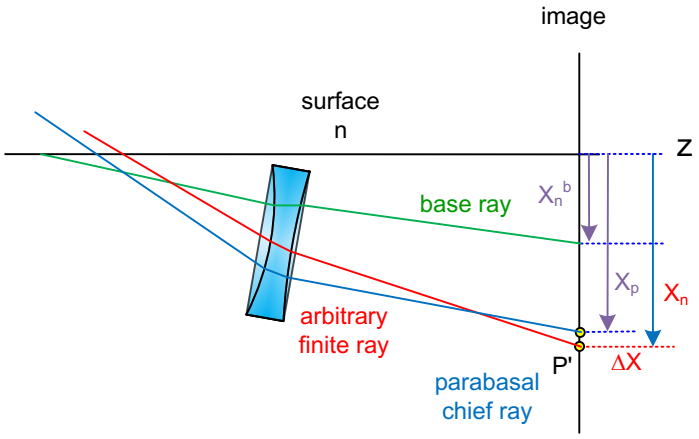



**Figure 8.** Illustration of the image side of the parabasal ray tracing.

The left hand side of Eq. (**45**) can then be written as

$$\frac{X_n \cdot n_n \cdot \overline{bx_n}}{\mu_n} - \frac{X_p \cdot n_n \cdot \overline{bx_n}}{\mu_n} + \frac{X_p \cdot n_n \cdot \overline{bx_n}}{\mu_n} - \frac{X_p \cdot n_0 \cdot \overline{bx_n}}{\mu_0} + P, \tag{47}$$

with $P = \left( X_n^b \cdot \overline{bx_n} - X_0^b \cdot \overline{bx_0} \right) \cdot \frac{n_0}{u_0}$.

Taking these back into Eq. (**45**), we obtain

$$\frac{\overline{bx_n} \cdot n_n}{\mu_n}\left( X_n - X_p \right) = \sum_{s=1}^{n} [X_s(\frac{\overline{bx_s} \cdot n_s}{\mu_s} - \frac{\overline{bx_{s-1}} \cdot n_{s-1}}{\mu_{s-1}}) - Z_s(\frac{\overline{bx_s} \cdot l_s}{\mu_s} - \frac{\overline{bx_{s-1}} \cdot l_{s-1}}{\mu_{s-1}}) \tag{48}$$
$$-\overline{\beta x_s}\left( l_s - l_{s-1} \right) - \overline{bx_n} \cdot X_p \cdot \left( \frac{n_s}{\mu_s} - \frac{n_{s-1}}{\mu_{s-1}} \right)] - P$$

The term P cannot be attributed to surfaces as their individual contribution; it therefore serves as a correction term for the decentered image plane. If the object plane is set as the reference for the global coordinate system, the condition $X_0^b = 0$ is fulfilled. Only when the base ray also hits the image plane at height zero, the quantity P automatically falls to zero, otherwise its magnitude remains proportional to $X_n^b$, which can be very large for many 3D systems.

The expression in y-z plane is written similarly as

$$\frac{\overline{b_n} \cdot n_n}{\mu_n}\left( Y_n - Y_p \right) = \sum_{s=1}^{n} [Y_s(\frac{\overline{b_s} \cdot n_s}{\mu_s} - \frac{\overline{b_{s-1}} \cdot n_{s-1}}{\mu_{s-1}}) - Z_s(\frac{\overline{b_s} \cdot m_s}{\mu_s} - \frac{\overline{b_{s-1}} \cdot m_{s-1}}{\mu_{s-1}}) \tag{49}$$
$$-\overline{\beta_s}\left( m_s - m_{s-1} \right) - \overline{b_n} \cdot Y_p \cdot \left( \frac{n_s}{\mu_s} - \frac{n_{s-1}}{\mu_{s-1}} \right)] - Q.$$

with $Q = \left( Y_n^b \cdot \overline{b_n} - Y_0^b \cdot \overline{b_0} \right) \cdot \frac{n_0}{u_0}$, $Y_p$ being the Guassian image height in y in the global coordinate system.

$$\Delta X = X_n - X_p = \frac{\mu_n}{\overline{bx_n} \cdot n_n} \sum_{s=1}^{n} [X_s(\frac{\overline{bx_s} \cdot n_s}{\mu_s} - \frac{\overline{bx_{s-1}} \cdot n_{s-1}}{\mu_{s-1}}) - Z_s(\frac{\overline{bx_s} \cdot l_s}{\mu_s} - \frac{\overline{bx_{s-1}} \cdot l_{s-1}}{\mu_{s-1}})$$
$$-\overline{\beta x_s}\left(l_s - l_{s-1}\right) - \overline{bx_n} \cdot X_p \cdot \left(\frac{n_s}{\mu_s} - \frac{n_{s-1}}{\mu_{s-1}}\right)] - P' \quad \textbf{(50)}$$

$$\Delta Y = Y_n - Y_p = \frac{\mu_n}{\overline{b_n} \cdot n_n} \sum_{s=1}^{n} [Y_s(\frac{\overline{b_s} \cdot n_s}{\mu_s} - \frac{\overline{b_{s-1}} \cdot n_{s-1}}{\mu_{s-1}}) - Z_s(\frac{\overline{b_s} \cdot m_s}{\mu_s} - \frac{\overline{b_{s-1}} \cdot m_{s-1}}{\mu_{s-1}})$$
$$-\overline{\beta_s}\left(m_s - m_{s-1}\right) - \overline{b_n} \cdot Y_p \cdot \left(\frac{n_s}{\mu_s} - \frac{n_{s-1}}{\mu_{s-1}}\right)] - Q' \quad \textbf{(51)}$$

The final expression for transverse aberrations for 3D systems are Eq. (**50**) and Eq. (**51**), which look similar to the original expression with additional correction terms $P' = \left(X_n^b \cdot \overline{bx_n} - X_0^b \cdot \overline{bx_0}\right) \cdot \frac{n_0 \cdot \mu_n}{u_0 \cdot \overline{bx_n} \cdot n_n}$, and $Q' = \left(Y_n^b \cdot \overline{b_n} - Y_0^b \cdot \overline{b_0}\right) \cdot \frac{n_0 \cdot \mu_n}{u_0 \cdot \overline{b_n} \cdot n_n}$.

The hidden assumptions of the derivation of the Aldis theorem for more generalized system geometries are $\overline{\beta x_0} = \overline{\beta x_n} = 0$ (or $\overline{\beta_0} = \overline{\beta_n} = 0$ for the y-z plane) and $Z_n = Z_0 = 0$ according to section 3.1. The first relation is always valid for the Gaussian image plane n, while the second relation is only true to perpendicular object and image planes. This assumption limits further generalization of the Aldis theorem for systems with tilted or even curved object or image planes. When the second relation fails to be true, the remaining terms can be moved to the right hand side of Eq. (**45**), becoming a part of the correction term.

## 3.5 EXAMPLES

Two selected examples are illustrated in this section to indicate the possible application of the extended Aldis theorem. The first anamorphic system example has a simpler on-axis layout, where the extension of the Aldis theorem to freeform surface shape is demonstrated. The second Alvarez lens example has decentered elements, where the method of parabasal ray tracing is applied.

The example anamorphic system shown in Figure 9 has two biconic surfaces placed at surface 3 and surface 5. With system stop located at surface 9, the system has an

image side f/# of 1.85 for both x and y. The maximum angle field in x is 12°, which is only half of that of y field. With an anamorphic factor of 2, the system is able to form a squared image on the right hand side.

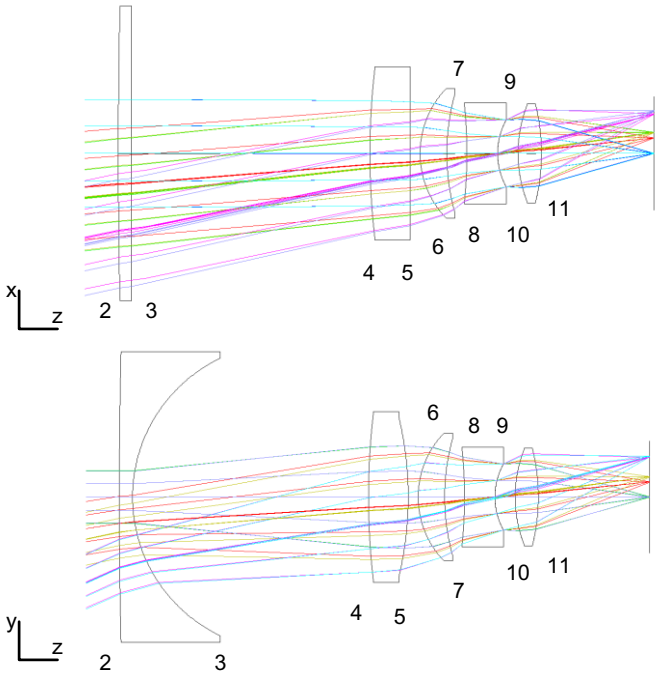
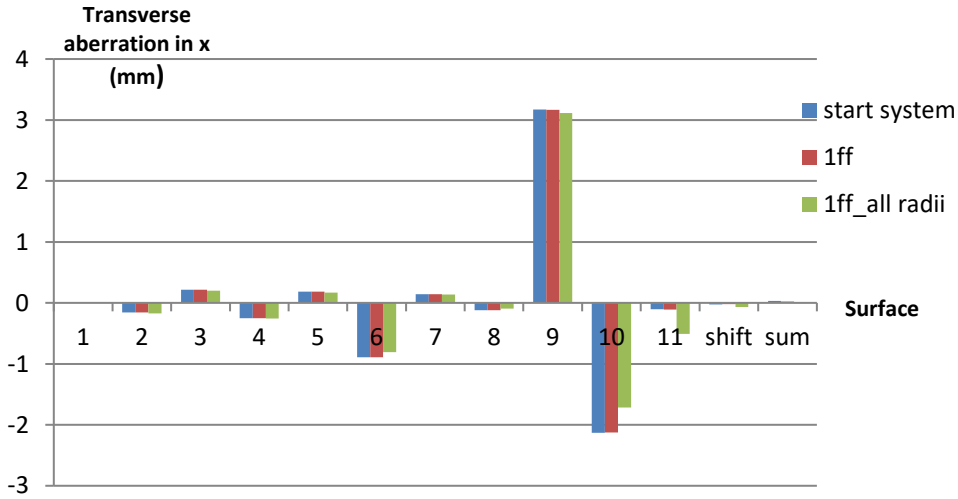



**Figure 9.** Layout of the anamorphic projector system.

The surface contributions for two different rays are shown in Figure 10 for three systems. The start system is the system without freeforms as shown in Figure 9. The 1ff system has one freeform at the last surface 11, and only parameters of the freeform are used for optimization. The last system is similar to the 1ff system, but with all surface radii included for optimization. The comparison between the first two system types shows that only the surface contribution of the freeform, the last surface 11, is affected when merely the freeform is optimized. With all surfaces considered for optimization, all surface aberration contributions change. The overall aberration of the system is decreasing with three systems, showing the performance improvement brought by the freeform.


(a)




(b)

**Figure 10.** Surface aberration contribution comparison between three systems for (a) ray (0,1,0.7,0.7) and (b) ray (1,0,0.7,0.7). The start system is the original system with two biconic surfaces. The 1ff system is the system with one freeform placed at surface 11, only the freeform parameters are optimized. The last system is similar to the second system, except that all surface radii are optimized.

Another example is an Alvarez lens system with real focusing lenses as shown in Figure 11. The focusing doublet placed after the Alvarez lens pair has a focal length of 80mm. This system configuration has a focal length of 40mm achieved by the decentration of both Alvarez lenses by 0.8mm towards opposite directions. The system stop is located in front of the Alvarez lenses at surface 1. In this example, real lenses that also contribute to the overall aberrations are introduced to show the

zooming effects of the Alvarez lenses. The reason for that is ideal lenses cannot be identified properly by the Aldis theorem. In the assumption of the Aldis theorem, ray directions only change when refractive index after the surface is not equal to 1, this surface will automatically contribute aberrations due to the ray bending. Nevertheless, ray bending still exists for the artificially defined ideal lenses, whose material cannot be specified. This unphysical model cannot be handled by the theorem; therefore surface contributions of ideal lenses given by the Aldis theorem are no longer correct.

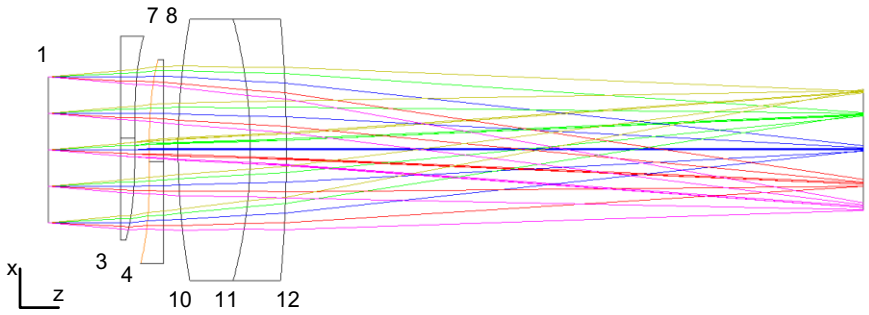



**Figure 11.** Layout of the Alvarez lens system with real focusing lenses. The configuration shown above has a focal length of 40mm.

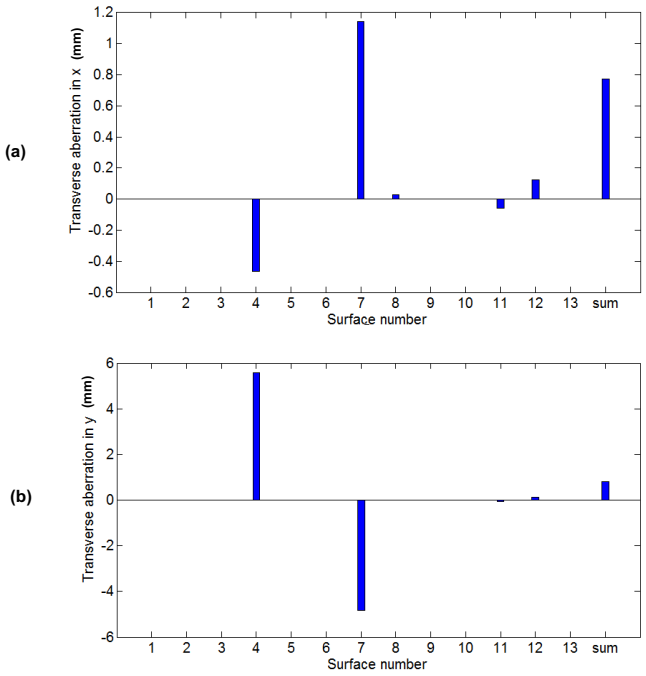



**Figure 12.** Alvarez lens system with real focusing lenses, surface contribution of the transverse aberration (a) in x and (b) in y for the field (1,0) using PV method excluding the correction term P.

For the field (1 0), the surface contributions of the transverse aberration in x and in y are presented in Figure 12 using the PV method explained in section 3.3. The correction term P' has no impact on individual surface contribution, and is therefore not drawn on the graph. The two freeforms of the Alvarez lenses, surface 4 and surface 7, have the largest contributions as expected. But the compensating effect of these two freeforms is also obvious due to their similar freeform topology. Similar to the Seidel diagram, we are able to analyze the Alvarez lens system's sensitivity roughly from the surface aberration contribution. The two freeforms, without doubt, will be the more sensitive elements during tolerancing.

Using the same sampling of 41×41, individual surface contributions of these two freeforms are given in Figure 13 (a) (b) for the complete pupil for the transverse aberration in x of the field (1, 0). Not only the compensating effects of both freeforms are indicated by the opposite signs over the whole pupil, but the shift of the central point due to element movements is also seen for both freeforms. Since both Alvarez lenses are shifted along the x-direction, the transverse aberration in y still has the minimum aberration point in the center as shown in Figure 13 (c) (d).

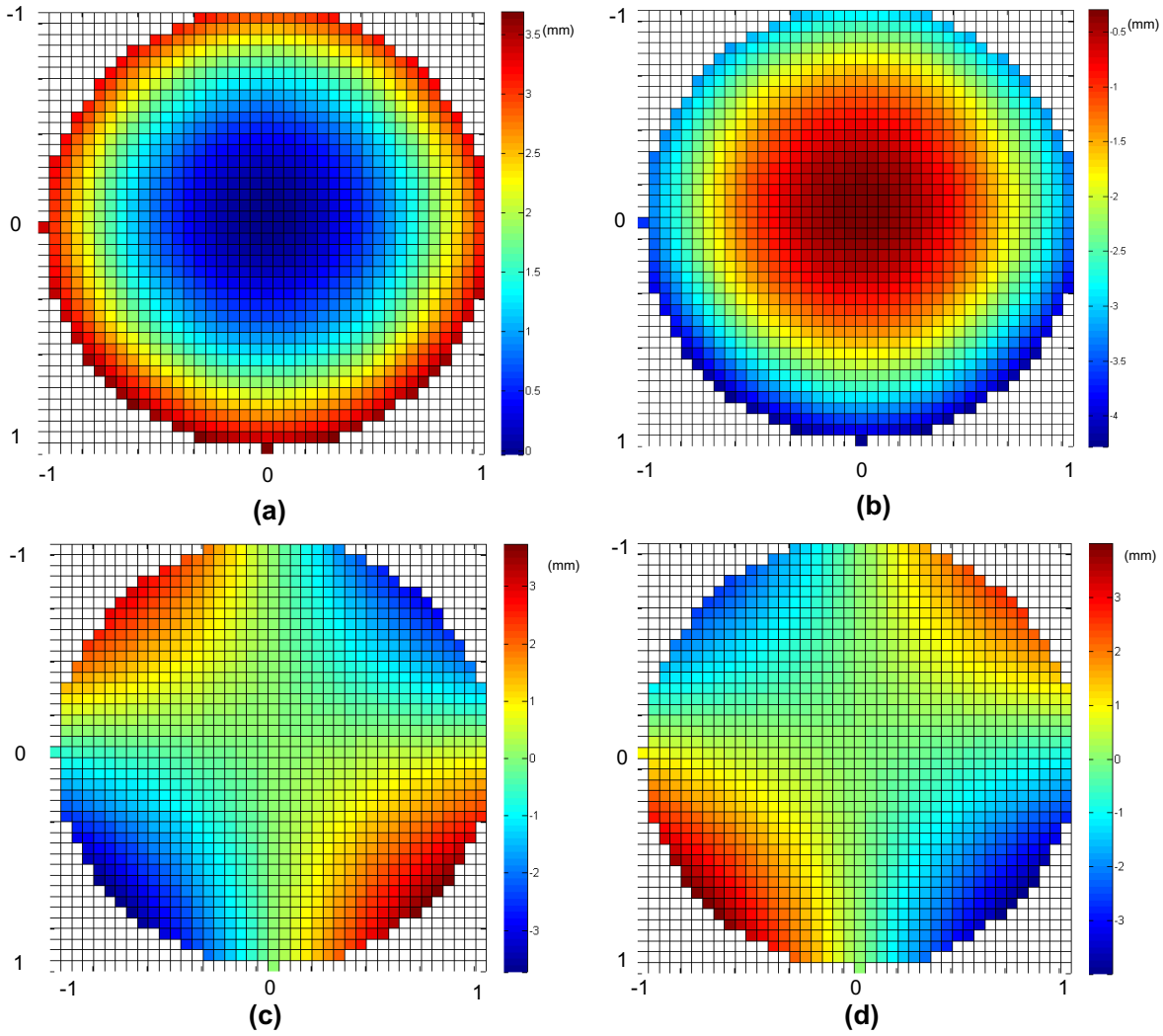



**Figure 13.** Full pupil distribution of the transverse aberration contribution for the field (1 0) (a) in x for surface 4, (b) in x for surface 7, (c) in y for surface 4, (d) in y for surface 7.

# 4 Optimization strategy with freeform surfaces

Progress in optical manufacturing and metrology [1, 2] opens widely the door for using freeform surfaces in optical systems [62, 63]. While the ever-growing demand on system features translates to increasing system complexity, the limitations of the traditional rotationally symmetric surfaces are clearly seen. Rotationally asymmetric freeform surfaces, on the other hand, with the decoupling in tangential and sagittal planes, provide more degrees of freeform for aberration correction. Nonetheless, the implementation of freeform surfaces is still at the investigation stage lacking systematic workflow. The optimal freeform locations inside the setup, as well as optimization strategies for freeform implementation, will be discussed in detail in this chapter. The numerical and empirical findings summarized here can serve as the first guidance for freeform surface optimization in imaging systems.

Two systems [64] are introduced here as investigation examples for this chapter. The Scheimpflug system shown in Figure 14 has an f/3 and is composed of twenty-two spherical surfaces, with one plane stop at surface 6. The tilts of the object and image planes lead to the broken symmetry of this system. The system suffers from a large spherical aberration due to the large numerical aperture (NA), as well as large coma and astigmatism. The inherit large keystone distortion of the system is quasi-paraxial, therefore can barely be corrected, the weightings of the distortion correction operands in the merit function are adjusted so that the main contribution comes

from the resolution correction. The RMS spot radius has been taken as the criterion for system resolution optimization in the merit function.

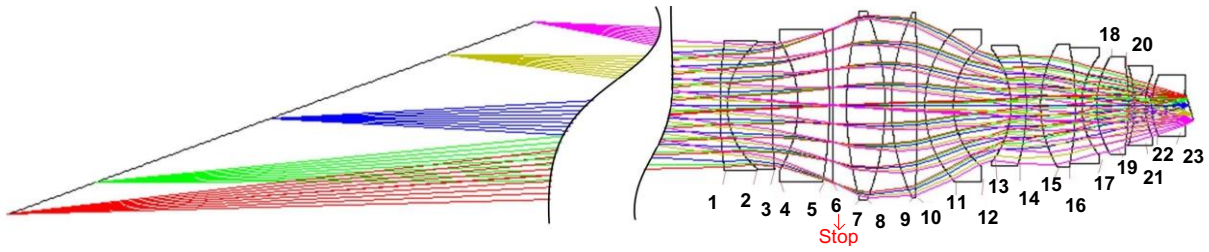



**Figure 14.** Layout of the Scheimpflug system [64].

The lithography system shown in Figure 15 has an f/0.43 with water immersion. The system has forty spherical surfaces and one plane stop. The object field is decentered to avoid the central obscuration of two mirrors. The lack of rotational symmetry due to the introduction of the mirror components makes the system a good candidate for freeform implementation. Extra access to the field plane is available due to two existing intermediate image planes between the mirrors. The original patent [65] employs twelve aspheres, whereas the system here is re-optimized with only spherical surfaces, and leaves plenty of room for improvement of dominating field aberrations. The merit function composed for the lithography system controls system parameters, telecentricity, distortion, and resolution, and provides a comprehensive evaluation of the system performance. Weightings of different operands are balanced so that no strongly dominant contributions occur.

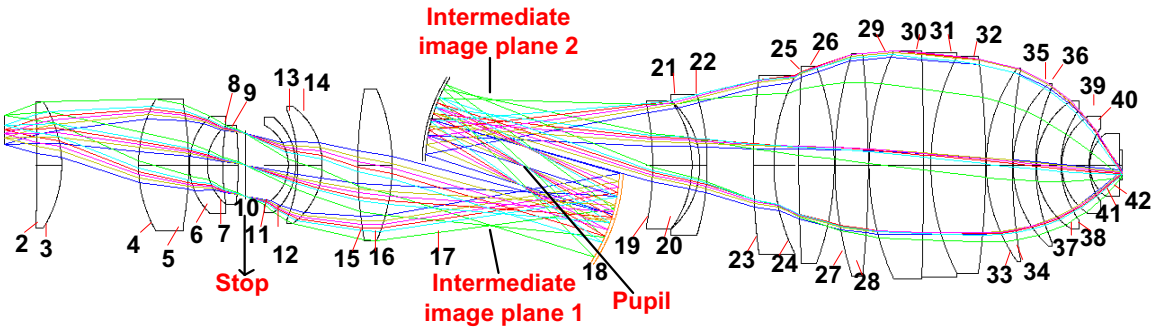



**Figure 15.** Layout of the lithography system [64].

The analysis of both start systems is used for constructing appropriate merit functions, which are conserved during the whole investigation process. As Zemax is already selected as the optical design software, unless mentioned specifically, the investigation in this chapter uses the freeform surface type "Zernike fringe sag", which is already implemented in the software.

## 4.1 OPTIMIZATION WITH ONE FREEFORM SURFACE

As will be shown in this section, the normalization radius of the freeform surface, the surface eccentricity, and the freeform order all have large impacts on the freeform implementation.

### 4.1.1 IMPACT OF THE NORMALIZATION RADIUS

The normalization radius of the Zernike fringe sag description is defined in Zemax separately from the surface clear aperture. If we assume the ratio between the normalization radius and the surface radius to be k, as illustrated in Eq. (**52**), the chosen value of k will have an impact on the system aberration correction.

$$k = \frac{norm.\_radius}{surface\_radius} \tag{52}$$

If k<1, the normalization radius is smaller than the surface radius, meaning the freeform is defined over a smaller area. If polynomial extrapolation for the surface area outside the normalization radius is not allowed, as indicated in Figure 16, a step-like discontinuity appears mainly due to the fact that freeform deformation beyond the normalization radius is not allowed. Since the area outside the defined freeform is still a sphere, no high order corrections are possible for the outer ring rays. Extrapolation should, therefore, be used in any case for proper correction of the complete ray set. If one allows for extrapolation, the edge of the freeforms may not be controlled properly as shown in Figure 17. The difference is quite notable as demonstrated in Figure 16 where freeform surface sag is not allowed for

extrapolation. The edge sag is only 0.26mm, in contrast to freeform surface sag with extrapolation in Figure 17, but causing a large sag value of 17.6mm.

If k>1, the normalization radius is larger than the surface radius, meaning only the central area of the defined freeform contributes to the system performance. In the case of low order Zernike polynomials, the central part is relatively flatter compared with the edge if k>>1. In consequence, the freeform will only weakly contribute to the overall correction. For such a scenario, the extension to a higher order is mandatory for improving the system performance.

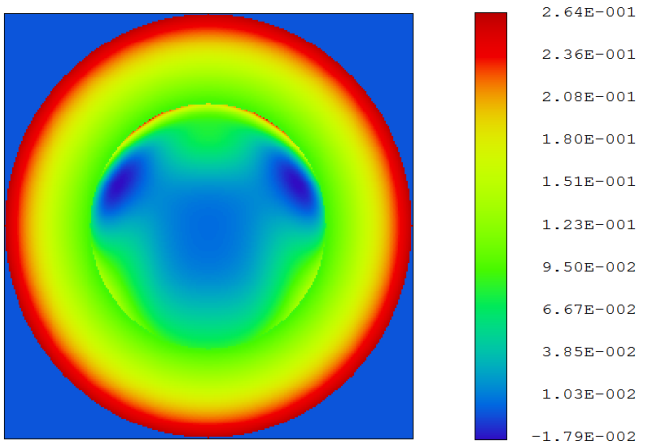



**Figure 16.** Freeform surface sag, k<1, no extrapolation. The surface sag is not continuous with edge sag of 0.26mm [64].

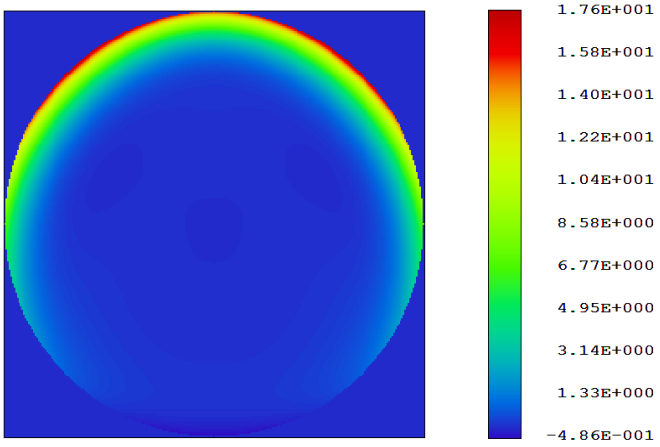



**Figure 17.** Freeform surface sag, k<1, with extrapolation. The freeform has large edge sag of 17.6mm [64].

Aiming for an optimal one-freeform position, a workflow based on Zemax macro is implemented, as shown in Figure 18. This macro allows comparing the resulting performance when placing the freeforms at different surface locations. System symmetry is additionally taken into account. For example, only x-symmetric Zernike polynomials are used for optimization, if the system has a conserved symmetry in the x-z plane. This overcomes the fact that Zemax is not able to detect system symmetry and will use all preset variables for optimization. Within the macro, only variables of the freeforms and the repositioning of the image plane are considered to avoid the influences of other surfaces. Using this workflow, the complexity of the solution space remains reasonable and results in a much faster optimization.

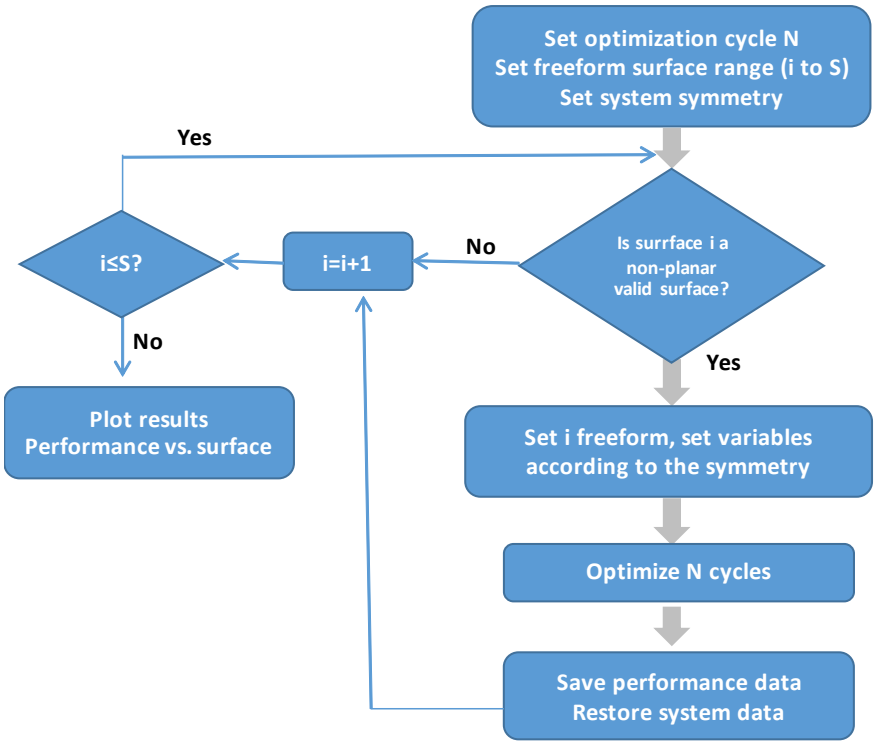



**Figure 18.** Workflow of the Zemax macro for determining the optimal one-freeform location [69].

The impact of the normalization radii with decreasing Zernike orders for the overall system performance is indicated in Figure 19 for the Scheimpflug system. In this diagram, freeforms with different Zernike orders are placed in different surface locations to study particularly the impact of varying the parameter k, from k=0.5 to k=4. The differences between different k curves are the smallest with the highest

freeform order. The differences increase with the decrease of order, which agrees with the discussion above.

To achieve the best performance of the freeform, k has to be set close to 1. We choose k=1.2 here for the following studies.

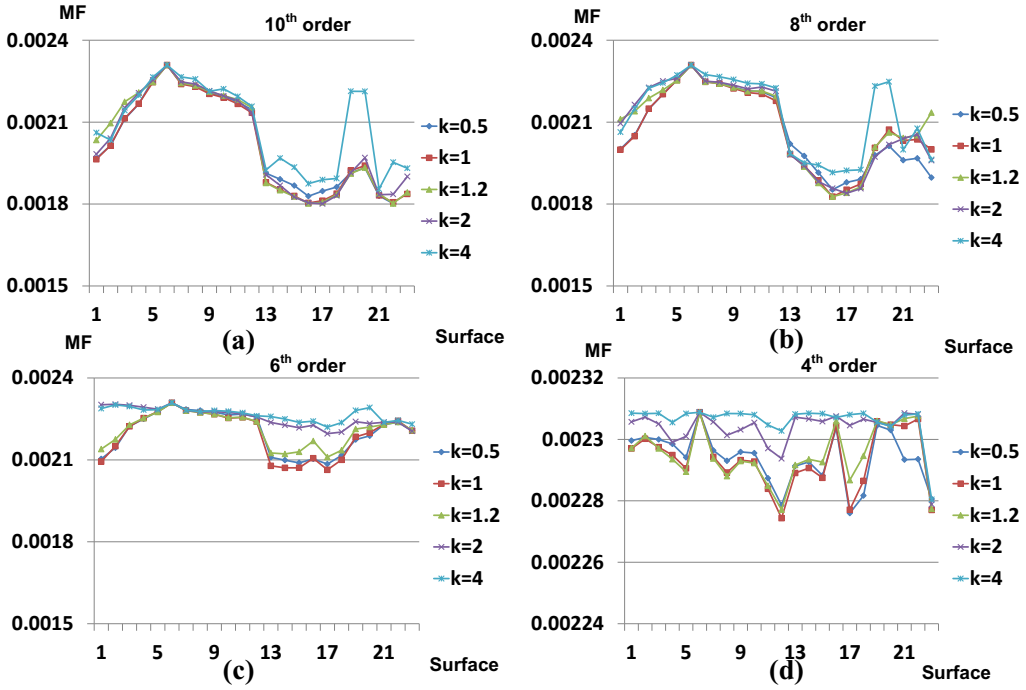


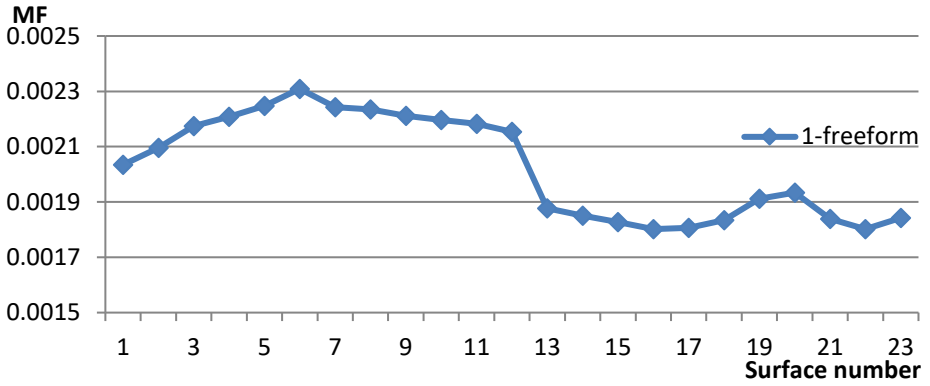

**Figure 19.** Scheimpflug system, impact of the normalization radii with decreasing Zernike orders. (a) freeform up to the 10th order, (b) freeform up to the 8th order, (c) freeform up to the 6th order, (d) freeform up to the 4th order. The y-axis corresponds to the merit function value, whose scaling is identical except for (d). The x-axis corresponds to freeform surface location [64].

### 4.1.2 SURFACE ECCENTRICITY

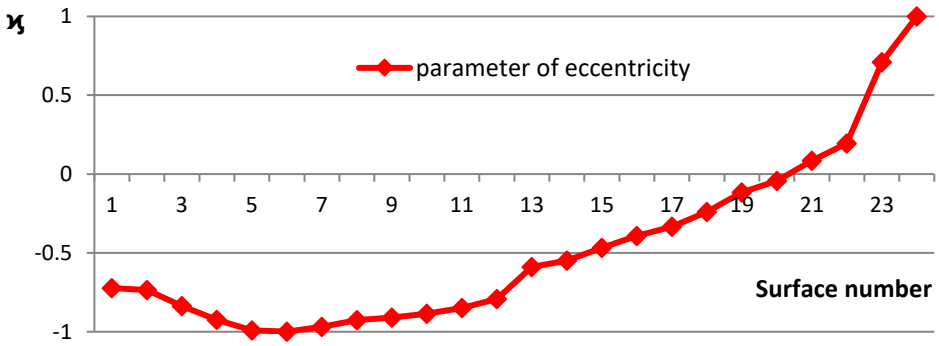
The asymmetry of the off-axis ray bundles results in field aberrations, this broken symmetry is where freeforms are especially beneficial. Intuitively, freeforms are better placed at locations where the ray bundles from different field points are well separated. To support this claim, we introduce a parameter called eccentricity, which describes the relative location of the surface with respect to the stop and field locations in a normalized form, i.e. $-1 \le \chi \le 1$. $h_{CR}$ represents the absolute height of the chief ray, and $h_{MR}$ represents the absolute height of the marginal ray. Obviously $\chi = 1$ corresponds to the image plane, whereas $\chi = -1$ corresponds to the pupil plane.

$$\chi = \frac{h_{CR} - h_{MR}}{h_{CR} + h_{MR}} \qquad (53)$$

For the two investigated systems, only fields in y direction are of interest due to the broken symmetry in this direction. Therefore, the chief ray heights and marginal ray heights in y are used for the calculation of eccentricity.




(a)

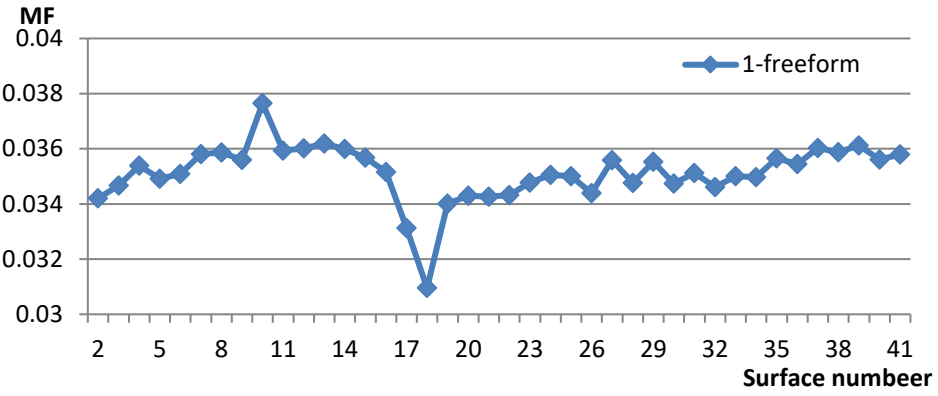



(b)

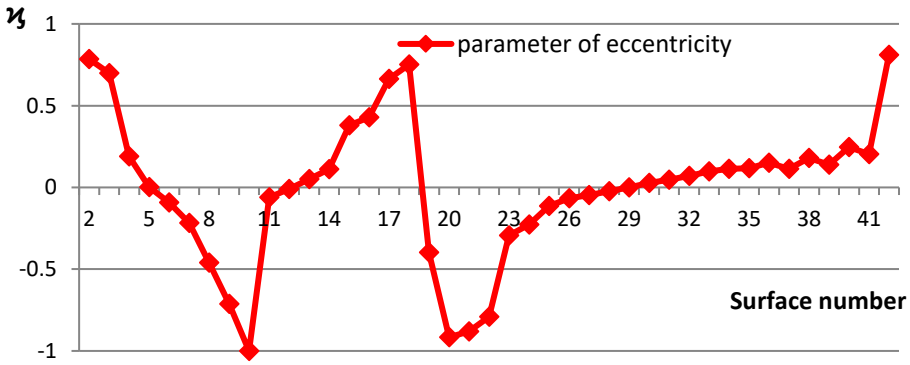
**Figure 20.** Scheimpflug system, one freeform performance vs. eccentricity. The x-axis represents the freeform surface number; the y-axis represents (a) the value of the merit function for the one-freeform curve (b) the eccentricity for the eccentricity curve [64].

In Figure 20, we relate the merit function values of the Scheimpflug system to the eccentricity, where the system stop is placed at surface 6. In Figure 21, this relation is extended to the lithography system. As shown earlier, the lithography system has two intermediate image planes between surface 17 and 18, and one stop placed at

surface 10. The extrema at surface 10 and surface 18 in the eccentricity diagram correspond to the extrema of the merit function values.




(a)




(b)

**Figure 21.** Lithography system, one freeform performance vs. eccentricity. The x-axis represents the freeform surface number; the y-axis represents (a) the value of the merit function for the one-freeform curve (b) the eccentricity for the eccentricity curve [64].

The Scheimpflug system, as well as the lithography system, has large field-dependent aberrations, which are even more pronounced in the latter. However, as a general observation, it is claimed that freeform locations away from the stop with larger chief ray heights are preferential. The spherical aberrations present at the Scheimpflug system are easier to compensate if the freeform is located near the stop location. The compromise between these two statements is related to the turning of the freeform performance curve in the rear part of Figure 20 (a). With the observation of an increase in the merit function values and a decrease in the

eccentricity parameter, we can easily evaluate the preferential locations of a single freeform. This observation has been helpful for the initial location of the freeform; nevertheless, the final freeform performance can only be optimized within a full design study. Using the knowledge of the system symmetry and the restricted number of chosen variables, the optimization speed of the macro improves drastically.

### 4.1.3 FREEFORM TOPOLOGY

Zemax allows describing the freeform topology via Zernike fringe sag surfaces up to the 10th order. This limitation can be easily expanded by additional user-defined Zernike polynomials. However, the question remains about the upper bound of the polynomial order. This has quite some impact on the manufacturing and feasibility of freeform optical components. Optimized system performance should be achieved with a possibly simpler freeform, which in many cases goes in parallel with a low order of the Zernike polynomials. Lower orders of freeform are in general more preferred in terms of optical manufacturing. On the other hand, one of the main advantages of the freeform surface is its ability to correct high order aberrations with its high order terms. For systems with higher even dominating high order aberrations, freeforms of at least the equal order are necessary for a good correction. In addition, adding high order terms translates into more run time for the optical design, while contributions coming from the higher orders are becoming smaller. In summary, it is neither convenient nor necessary to use too high order polynomials when introducing freeforms for designing and optimizing an optical system.

For our chosen Scheimpflug system, the main aberrations include the Zernike terms Z11, Z12, Z15 and Z21, which belong to the $6^{th}$ and the $8^{th}$ order. The impact of the polynomial order for the Scheimpflug system performance is indicated in Figure 22, where freeform performance curves with different orders are represented by different colors. Z21, out of four main aberration terms, shows the smallest contribution. The system's aberrations could in principle be well compensated by 8th order Zernike polynomials. $4^{th}$ order freeforms, on the other hand, lack the

ability for high order aberration correction. At this stage of the design, the best result is achieved by locating the freeform at surface 16.

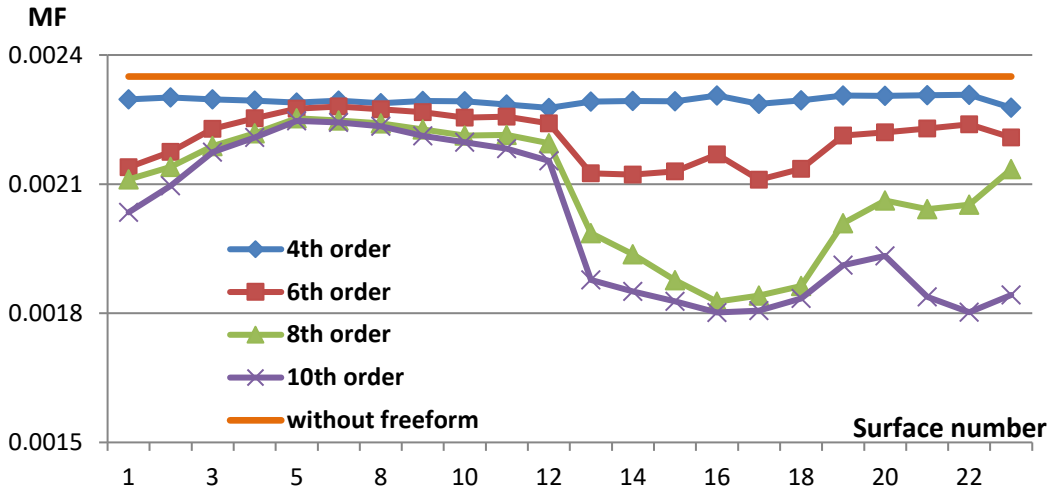



**Figure 22.** Scheimpflug system, impact of the freeform polynomial order. The y-axis represents the merit function value; the x-axis represents the freeform surface number. Colored curves correspond to one-freeform performance using freeforms up to the respective orders. The orange straight line indicates performance of the original system. An improvement factor of 1.31 is achieved for the optimal surface location [64].

For the lithography system, the main aberrations are Z5 and Z11, respectively Z9 and Z17, and to a less pronounced extent Z12, Z20 and Z27. The large aberration Z12 and Z11 cannot be corrected by a 4$^{th}$ order freeform as shown in the blue curve in Figure 23. While 6$^{th}$ order aberrations are well compensated by the 6$^{th}$ order freeform surfaces, the remaining 8$^{th}$ order Z17 needs higher order freeforms. For a fast evaluation and determination of the optimal freeform location, the Zernike polynomials up to the 8$^{th}$ order prove to be sufficient. At this stage of the design, we decide the best result is achieved by locating the freeform at surface 18.

To be noted, knowledge of the aberration theory up to the 8$^{th}$ order is necessary here in order to solve the considered problems analytically, but the currently available formulas for nodal aberration theory [7-11] are only suitable for freeforms up to the 6$^{th}$ for the basic shape. It becomes extremely complicated to consider the overall resulting effects when adding high order terms [14]. Therefore, the analytical solution is neither suitable nor substantial for solving freeform positioning problems

for complex multi-lens systems. Nevertheless, surface-resolved contribution representation of higher orders [26] helps to understand the aberration compensation of a system, which may give the first indication of strategy for correction.

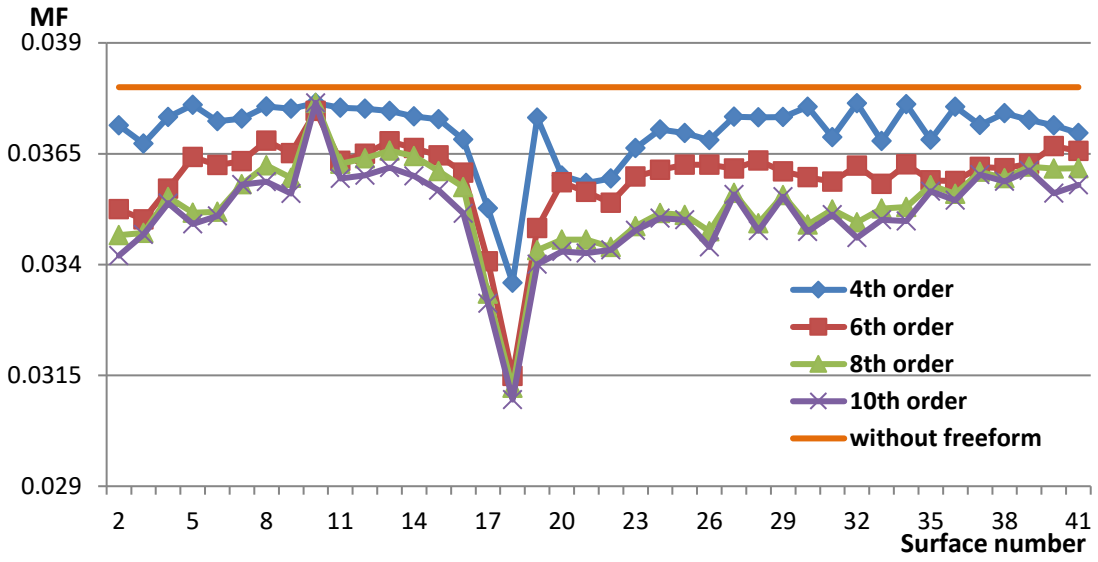



**Figure 23.** Lithography system, impact of freeform polynomial order. The y-axis represents the merit function value; the x-axis represents the freeform surface number. Colored curves correspond to one-freeform performance using freeforms up to the respective orders. The orange straight line indicates performance of the original system. An improvement factor of 1.23 is achieved for the optimal surface location [64].

## 4.2 Optimization with two freeform surfaces

For many systems, the design strategy disclosed in section 4.1 based on the implementation of a single freeform maybe sufficient to reach the required system specifications. Nevertheless, more degrees of freedoms become available when adding additional freeforms. However, this "more than one" approach translates immediately into a higher system complexity. Especially the interplay between multiple freeforms is worth to be studied in more detail. Based on the nodal aberration theory, it may be beneficial to place the first freeform at locations close to $\chi$ =-1 to achieve a field-independent aberration correction, and to place the additional freeform at locations with $\chi$ close to +1 to compensate the field-

dependent aberrations. These general statements are helpful for launching the initial design but will in no case replace a full design study.

To assist our investigation, the Alvarez lens system is introduced to help understand the interactions between two freeform surfaces. The Alvarez lens pair is the simplest two-freeform system, where two identical 6$^{th}$ order freeforms are placed together, whose vertical displacement contributes to the systems' focal length variation. As explained in section 2.4.2, induced aberration originated from the 6$^{th}$ order freeforms cannot be corrected with the classical setup. Freeforms with higher orders or more freeforms are necessary for better system performance. As shown in Figure 24 and Figure 25, a simple Alvarez lens system is constructed with two glass plates, whose inner surfaces (surface 4 and 7) are freeforms. One ideal lens is positioned at surface 10 to help focus the light beam. The system stop is placed prior to the Alvarez lenses. With maximum allowed lateral displacement $\Delta x$ of $\pm 0.5$mm for a 14mm-diameter lens [46], the dynamic range of the focal length lies between 80mm and 120mm. The achievable focusing range depends on both A and $\Delta x$. Having a large A is more comfortable from the mechanical movement point of view, with the disadvantage of increased induced aberrations. The decision of the maximum $\Delta x$ depends on the application of the system as well as the technology used for movement. The classical Alvarez lens pair has only an on-axis field. With the consideration of fields of $\pm 5°$ in x, the system is optimized with the original freeforms terms (6$^{th}$ order) and the tilted image plane to have the minimum overall spot size. Rays coming from different fields end up having different focal lengths due to different intersection points at the freeforms, therefore the tilt of the image plane is necessary to cancel out partially the defocusing effects [68]. The Alvarez lens system can serve as a simple example for indicating the interplay between two freeform surfaces.

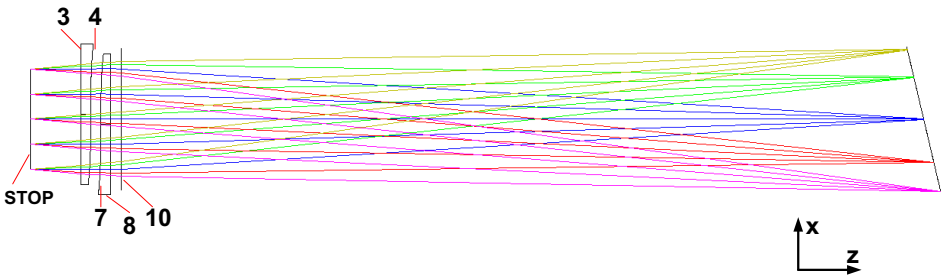



**Figure 24.** Layout of the Alvarez lens system with an ideal focusing lens, both on-axis and off-axis fields are drawn for one system configuration. The image plane is tilted to accommodate the defocusing effect.

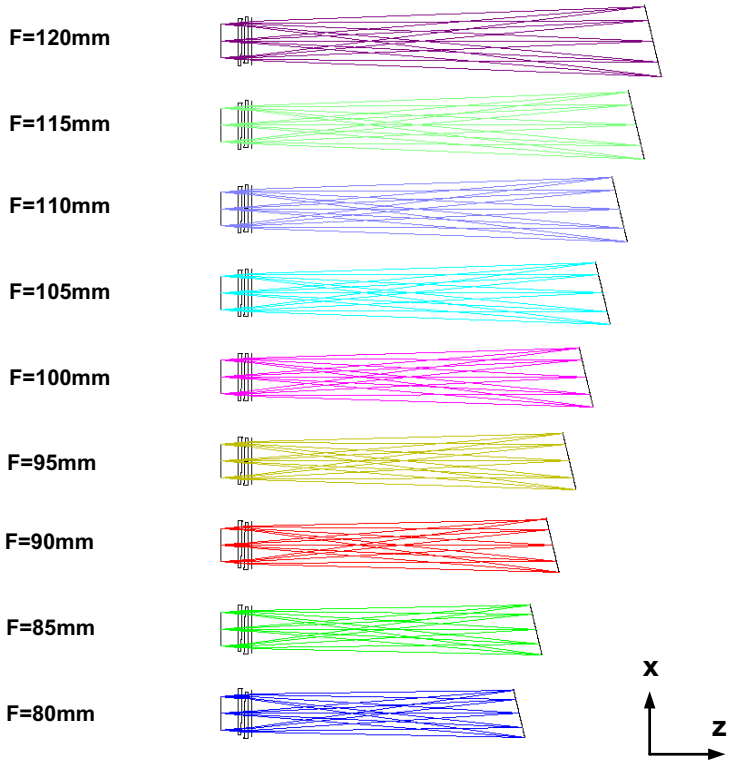



**Figure 25.** Layout of the Alvarez lens system with varying focal lengths.

### 4.2.1 OPTIMIZATION STRATEGY

Due to the complexity of the design, the considerations are limited to two freeforms being part of the complete story. The optimization strategy by itself leads to additional concerns about placing the freeforms in a successive approach or a simultaneous approach, and increasing the freeform polynomial orders successively

or simultaneously. Therefore, five distinct strategies are investigated to elucidate the search for the best optimization.

A. Two surfaces are optimized successively with fixed orders.
B. Two surfaces are optimized simultaneously with fixed orders.
C. Like strategy B, but the optimization cycles are done in smaller successive steps.
D. Two surfaces are optimized simultaneously with orders increased step by step.
E. Two surfaces are optimized successively with orders increased step by step.

For all strategies, I limit the variable set to the two freeform surfaces and the repositioning of the image plane in order to eliminate the impact of the residual surfaces. Strategy A and B are designed to investigate the impact of the optimization sequence. For strategy A, I add first the first freeform surface, then optimize it for 50 cycles before adding the second freeform surface, optimizing both surfaces together for another 75 cycles. For strategy B, two freeform surfaces are added and optimized together for100 cycles.

Optimization is a complex iterative process, which means the next iteration should always result in more improved system performance. Strategy C is similar to strategy B, but divides the large cycle of optimization into smaller steps with fewer optimization cycles. This has been an attempt to verify if strategy C may have an advantage over strategy B. Launching an optimization along strategy B against strategy C, comparable results between both strategies are expected.

Strategy D and E have been chosen to understand the difference between a simultaneous or successive introduction of Zernike orders. Strategy D intends to optimize two freeforms simultaneously starting with the lowest order. The Zernike orders are increased gradually for both surfaces simultaneously. Each step contains 25 cycles. In strategy E, the optimization starts with one freeform and the lowest Zernike order. While optimizing the second freeform with the lowest Zernike orders, all parameters of the first freeform stay fixed. Simply speaking, both freeforms are never optimized at the same time, and the introduction of higher orders is always

done in turns. The number of the cycle for each step is 25 so that the overall optimizations are still comparable with strategy A and B.

Strategy A and E are both optimizing two freeforms successively. The difference between both is, the freeform orders are increased to the maximum gradually for E, while staying fixed from the beginning at the $10^{th}$ order for A. Due to the limited number of variables, as well as the choice of the freeform surface representation, the optimization process converges quite fast for all strategies. The selection of the optimization cycle is sufficient for this investigation. Strategy B requires more optimization cycles to converge due to its more complex solution space. The comparison between B and A would lack confidence if an insufficient number of optimization cycles is taken. It is noteworthy that for a successful optimization comparison, the same merit function is maintained.

All optimization strategies are run on the Alvarez system and the results are summarized in Table 2. The merit function value, system resolution represented by the RMS spot radius, as well as freeform surface sag ranges are documented for comparison. It is noteworthy that the freeform sag here refers to the PV sag value of the freeform deformation part as shown in Eq. (9), which is somewhat different from the definition in manufacturing [38] in section 2.3.3.

As can be seen from Table 2, strategy A, B, C, D show no notable differences in terms of system performance and freeform sags, whereas strategy E exhibits significant worse optimization performance.

**Table 2.** Comparison of different optimization strategies using the Alvarez lens system. S4 represents the first freeform surface, and S7 represents the second freeform surface.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| **MF** | 0.002081586 | 0.002081599 | 0.002081619 | 0.002081592 | 0.003690935 |
| **RMS spot radius (mm)** | 0.002221 | 0.002221 | 0.002221 | 0.002222 | 0.003209 |
| **S4 sag (mm)** | -0.581~0.499 | -0.579~0.500 | -0.579~0.501 | -0.583~0.499 | -0.367~0.363 |
| **S7 sag (mm)** | -0.571~0.514 | -0.569~0.514 | -0.568~0.515 | -0.572~0.513 | -0.356~0.374 |

For the Scheimpflug system, the first freeform surface is fixed at surface 16 according to the guidance of one-freeform optimization. The second freeform, however, can be placed on any of the remaining surfaces. The system performance containing two freeforms represented by the merit function value is shown in Figure 26. We can see clearly similar observation as for the Alvarez lens system, that there are no notable differences among the first four strategies. It is supposed that the optimization is trapped in a local minimum with larger merit function values due to the limited number of variables of strategy E. Comparing the freeform PV sags of surface 16 and the second freeform, the first four strategies still show no significant differences as indicated in Figure 27 and Figure 28. Although some strategies give larger surface sags at some surface locations, the overall performance of individual strategy is comparable between strategy A, B, C and D.

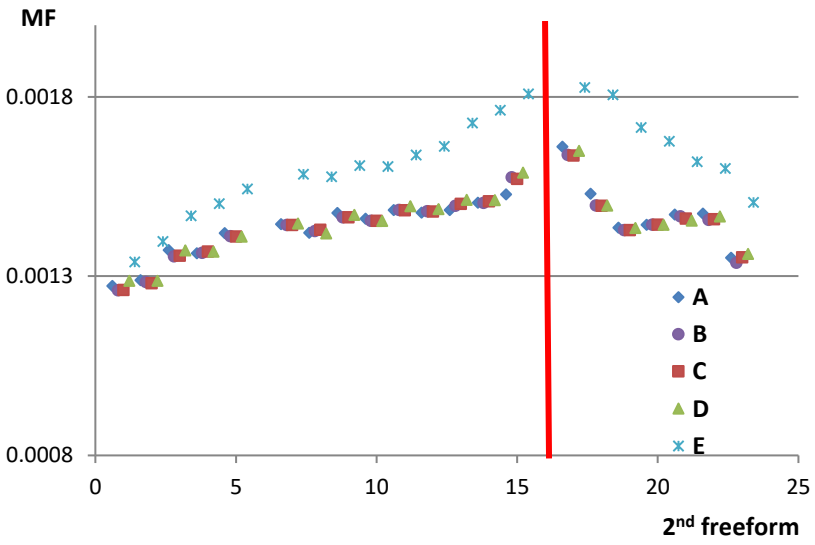



**Figure 26.** Scheimpflug system, optimization strategy. The first freeform surface is placed at surface 16, which is marked with a red line. The x-axis corresponds to the surface index of the second freeform. The y-axis is the value of the merit function as the criteria for performance comparison of different strategies. Results from different strategies are marked with different symbols [64].

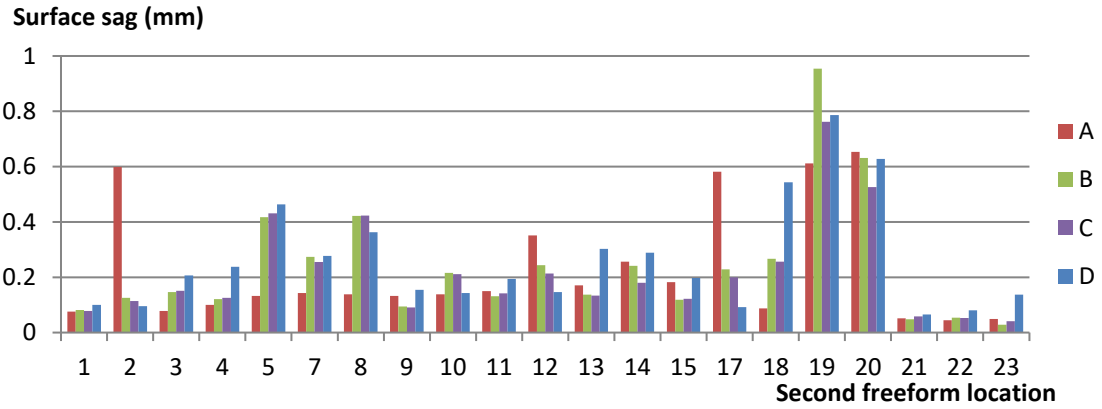



**Figure 27.** Freeform sags of surface 16 for different optimization strategies with respect to varied second freeform positions. The x-axis represents the second freeform location, the y-axis represents the freeform sag of surface 16.

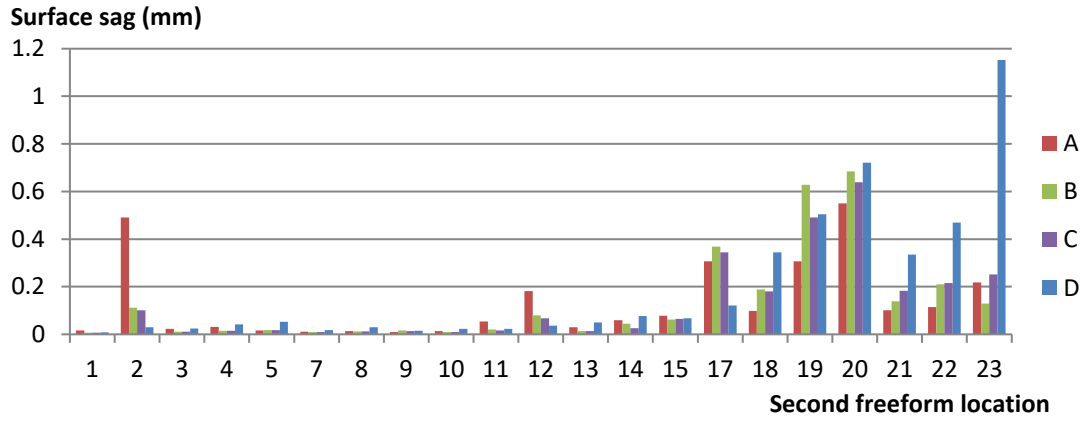



**Figure 28.** Freeform sags of the second freeform for different optimization strategies. The x-axis represents the second freeform location; the y-axis represents the freeform sag of the second freeform.

Similar observations are associated with the lithography system despite the more complex structure; the disadvantages of strategy E are also presented in Figure 29 in the front part of the surface locations. No clear choice of strategy A, B, C or D can be provided for the optimization of the lithography system; however, strategy C appears to be slightly better in comparison with strategy B at most locations. In the Scheimpflug example, the performance of the second freeform is significantly worse around the location of the first freeform. However, in the case of the lithography system, this finding is no longer true. The explanation for this is simple. Due to the existence of two intermediate image planes and one pupil plane between the mirrors, the change of surface eccentricity and surface behavior in between is almost invisible.

Our results indicate that freeforms can be optimized along strategy A, where the first freeform is fixed in location and is fully optimized before adding the next freeform surface for additional improvement of the system performance. The optimization can also be done in smaller steps to examine the system more frequently. In this way of implementation, fewer freeform surfaces are needed for the investigated systems without sacrificing optimized system performance.

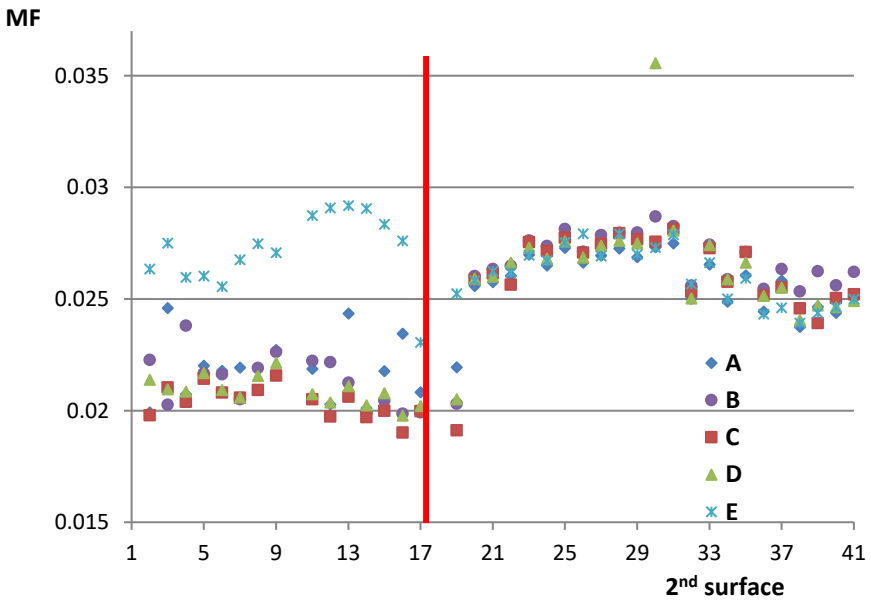



**Figure 29.** Lithography system, optimization strategy. The first freeform surface is placed at surface 18, represented by a red line. The y-axis is the value of the merit function, and the x-axis is the surface index of the second freeform. Results from different strategies are marked with different symbols [64].

### 4.2.2 FREEFORM INTERACTION

As shown in the previous section, the implementation of freeforms in a successive or a simultaneous manner shows no significant difference on both the system performance and the freeform sag value. Nevertheless, the interaction between two freeforms is still unknown. Especially the choice between one higher order freeform and an extra lower order freeform is not clear. Considering the complex freeform topologies of individual examples, it is hard to compare the potential costs of both possibilities. However, as a question that arises during system optimization, more

detailed investigation exploring the interplay between two freeforms is done in this section using the Alvarez lens system and the Scheimpflug system.

To explore the freeform interaction of the simple two-freeform system, the Alvarez lens system, the freeform orders are increased alternatively. In order to explore the solution space better, both freeforms are optimized together. The optimization strategy used here is the mixture of strategy D and E.

As indicated in Table 3, the freeform PV sags of the Alvarez lens system after optimization are documented, where a strong interaction between two freeforms is observed. Especially when both freeforms differ in orders, the freeform sags still remain in a comparable range for both surfaces. When the first freeform is increased to the 8$^{th}$ order and the second freeform remains at the 6$^{th}$ order, both freeforms undergo merely minor changes in PV sags. The same is true for the third scenario, where only the first freeform is increased to the 10$^{th}$ order. The differences between both freeform sags, on one hand, contribute for higher order aberration correction, on the other hand, is responsible for the varying focal lengths of the setup.

**Table 3.** Alvarez lens system freeform sags with alternatively increased freeform orders. The start system has a 6$^{th}$ order representation with a PV sag value of 0.73mm for the first freeform surface (surface 4).

| | **S4 8$^{th}$ order, S7 6$^{th}$ order** | **S4 8$^{th}$ order, S7 8$^{th}$ order** | **S4 10$^{th}$ order, S7 8$^{th}$ order** | **S4 10$^{th}$ order, S7 10$^{th}$ order** |
|---|---|---|---|---|
| **Surface 4 PV sag (mm)** | 0.73 | 1.206 | 1.222 | 1.231 |
| **Surface 7 PV sag (mm)** | 0.73 | 1.21 | 1.227 | 1.234 |

The freeform interactions are further explored with the Scheimpflug system. Using the same strategy, the orders of the chosen two freeforms, surface 16 and surface 1, are increased alternatively. In Figure 30, the rescaled merit function values, system resolution represented by the RMS spot radius (RSCE), and freeform sags are plotted. Unusually large freeform sags of the surface 16 start to appear from the 8$^{th}$ order, as pointed by the yellow arrow in the figure. To investigate the reason for this large sag, a different optimization approach is utilized in Figure 31, where the freeform orders of both surfaces are increased at one step from the basic spherical shape. We see clearly from Figure 31 that the large freeform sag of surface 16 only appears for one

scenario. Comparing Figure 30 and Figure 31, we can conclude the unusual behavior of the large surface sags in case of high order freeforms is related to the large freeform sag of one freeform combination. It appears that Zemax is not able to decrease surface sag (e.g. freeform coefficients) properly during optimization. The compensation of the system's residual aberrations is achieved more likely by continuously increasing freeform coefficients. In the case of freeform optimization, where extremely large degrees of freeform are introduced to the system, the current local optimization algorithm of Zemax seems to be less than ideal and may be insufficient for the optimal results. Direct connections between freeform sag values and freeform contributions cannot be established due to the drawback of the optimization algorithm.

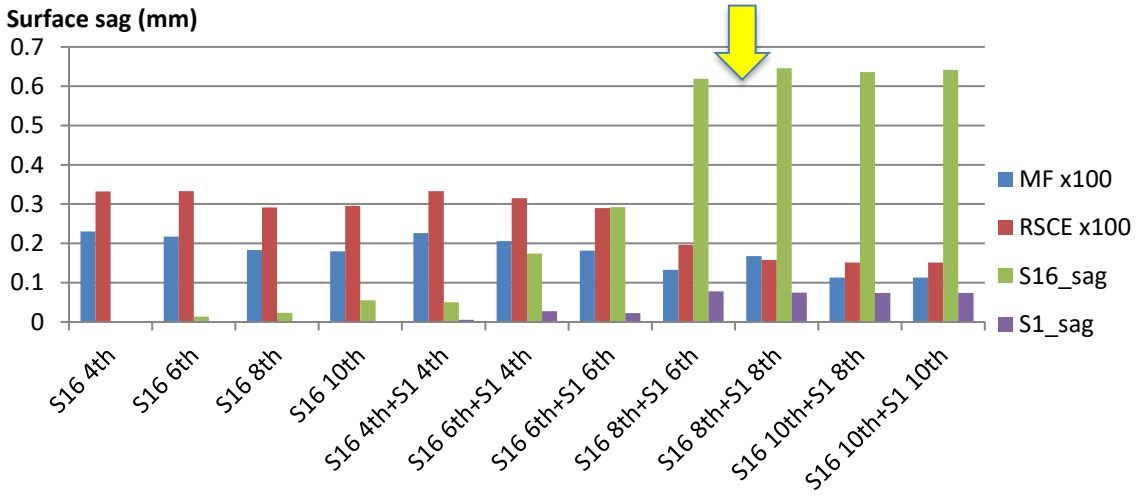



**Figure 30.** Scheimpflug system, freeform sags of surface 16 and surface 1 with increased freeform orders. The freeform orders are increased alternatively on the basis of the previous adjacent surface order combination.

Taking a closer look at Figure 31, we observe the fluctuation of system performance with respect to different freeform order combinations. For a single freeform surface 16, the system performance remains nearly constant when the freeform order increases from the $8^{th}$ to the $10^{th}$, which corresponds to what have already been shown in section 4.1.3, that only small contribution can be obtained if higher-than-necessary orders of the freeform are introduced.

Two 6th order freeforms give similar system performance as to that of one freeform up to the 8th order, however, the freeform sag values of the first case are significantly larger. Although the magnitude of the freeform sag cannot be translated directly into the magnitude of manufacturing difficulty, it still provides us insights when it comes to the trade-off between one higher order freeform and two lower order freeforms. In this case, one higher order freeform can be replaced by two lower order freeforms at the price of increased freeform sags. The second freeform should only be used when necessary, considering the manufacturing and assembly difficulties that come along with the extra freeform surface. Comparing surface 16 up to the 10th order with surface 16 up to the 8th order and surface 1 up to the 6th order, we notice immediately the significant decrease in the merit function value of the latter case. When further performance improvement is not significant with the increasing higher orders of the first freeform, the introduction of the second freeform is inevitable.

As the system performance improves with increased freeform orders, the freeform sags, on the other hand, is not continuously increasing. The interaction between two freeforms appears to exist, since a larger surface sag of the first freeform also corresponds to a larger surface sag of the second freeform.

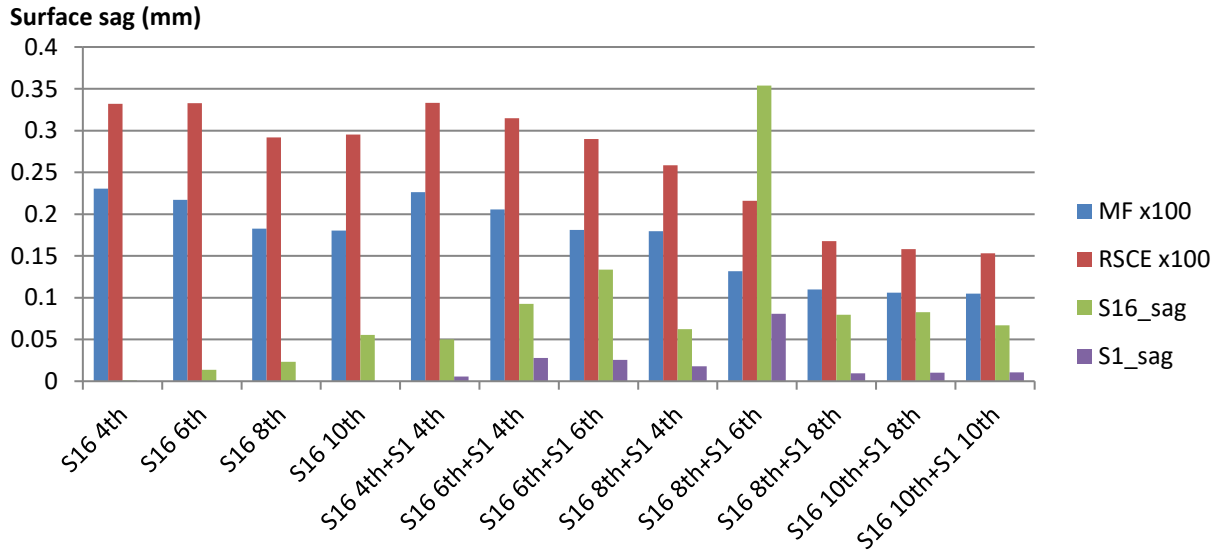



**Figure 31.** Scheimpflug system, freeform sags of surface 16 and surface 1 with increased freeform orders. The freeform orders are increased at one step from the basic spherical shape.

### 4.2.3 TWO FREEFORM MAPPING

Our results so far suggest that optimizing freeforms simultaneously or successively with fixed orders does not have a major impact on the final outcome despite the interplay between two freeforms. Therefore, it is suggested to add the freeforms successively. It is worth noting that system configuration is already changed after introducing the first freeform. As an example and as indicated in Figure 32 for the Scheimpflug system, the optimization results with the second freeform, indicated by the green path, place the first freeform at surface 3, whereas the purple and red paths place the first freeform at surface 16 and surface 20, respectively. Obviously, the system performance with the second freeform evaluated by the merit function behaves differently. Putting the second freeform close to the first freeform seems to be an inappropriate choice, in opposition to putting the freeforms in well-separated locations. As already mentioned, the system conditions change by the insertion of a freeform, demanding to run the aforementioned macro again prior to the insertion of the second freeform. We also notice that the relation between the merit function curves and the eccentricity curve is changed when compared with Figure 20. With the selection of different first-freeform positions, the overall performance of the system with two freeforms are also different. It appears better system performance is only obtained with an optimal first-freeform selection.

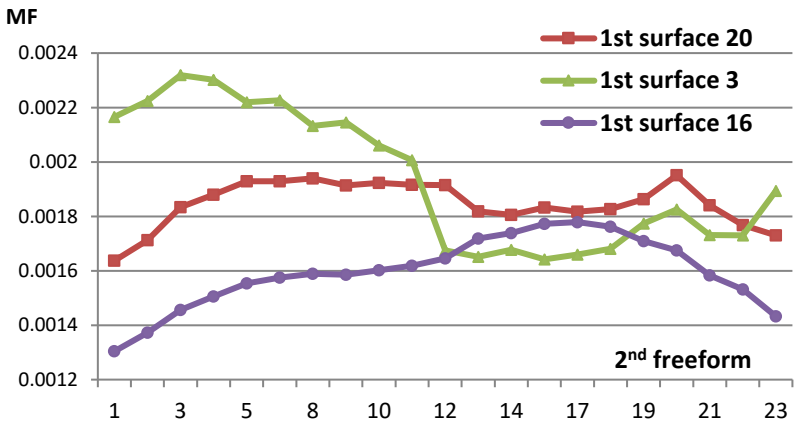



**Figure 32.** Scheimpflug system, system performance with a second freeform under different first freeform selections. The best freeform location changes after implementation of the first freeform surface [64].

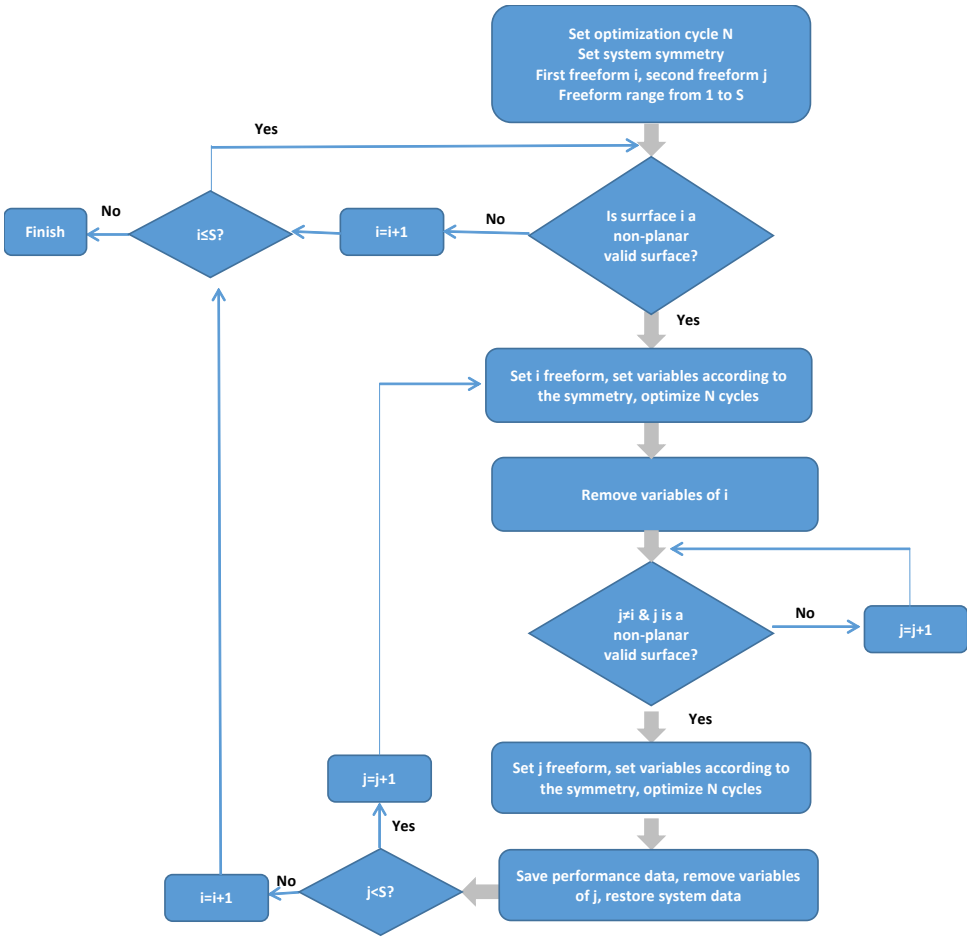



**Figure 33.** Workflow to implement all two-freeform combinations of an optical system [69].

I construct a macro for evaluating all possible two-freeform combinations, which is given in Figure 33. Every valid two-freeform combination is optimized using solely the freeform parameters and repositioning of the image plane for the optimal system performance. The optimization cycle is set to be fifty. Considering the amount of optimization needed, this process is much more time consuming than the one-freeform approach. The corresponding merit function values are shown in a two-freeform map as presented in Figure 34 for the Scheimpflug system. The two-freeform map provides on the y-axis the surface locations that have been used to implement the first freeform, whereas x-axis indicates the locations of the second freeform. The color coding of this map indicates the final merit function value. The white elements in this matrix are stop position at surface 6 and diagonal positions where both freeform positions coincide.

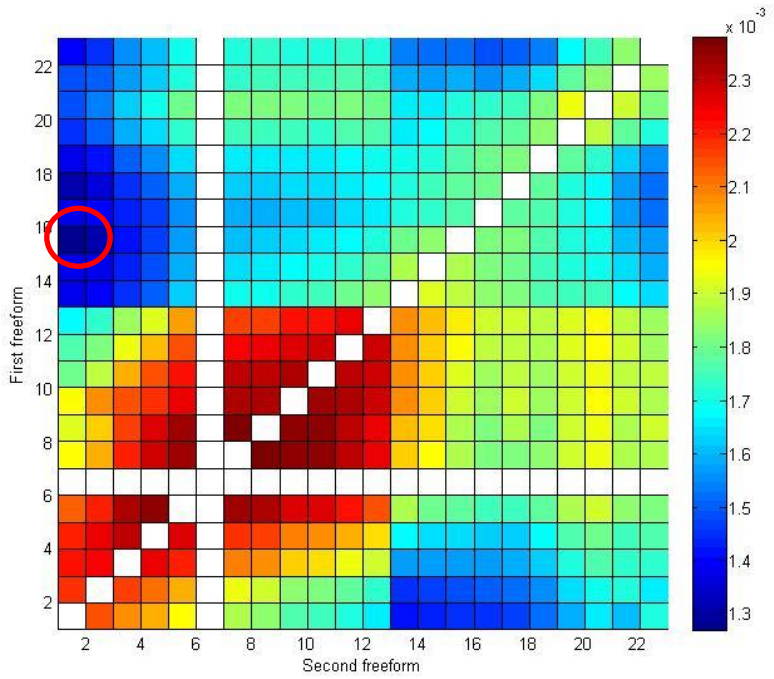



**Figure 34.** Scheimpflug system, two freeform mapping. The y-axis corresponds to the first freeform surface number, and the x-axis corresponds to the second freeform surface number. The white elements in the figure are invalid freeform surface locations. The red circle marks the area of optimal two-freeform combinations.

Figure 35 is obtained through $M_{ij}^{1} = M_{ij} - M_{ji}$ , where $M_{ij}$ represents the system performance described by the merit function value for the freeform choice i and j (i being the first freeform, j being the second freeform). Figure 35 indicates the trial to understand the performance differences when putting the first surface in position i and the second freeform in position j, or vice versa. The differences are almost a magnitude smaller. Figure 35 shows that putting the first freeform close to the stop position 6 and the second freeform close to position 14 appears to be a non-optimized sequence. The optimization process seems to be trapped in a worse local minimum mainly resulted by the bad choice of the location for freeform one. As already mentioned, it seems to be preferential to place freeform one at its optimal position first before placing the second freeform at other surface locations.

The optimal choice for freeform placements is indicated in Figure 34 for the Scheimpflug system by a red circle. This corresponds to the minima of the merit function values in Figure 20(a) and Figure 32(purple path). As indicated by Figure 34,

placing the two freeforms close by (close to the diagonal of the freeform map) is non-optimal, which corresponds to our first statement to place the freeforms well apart from each other when launching the second freeform optimization.

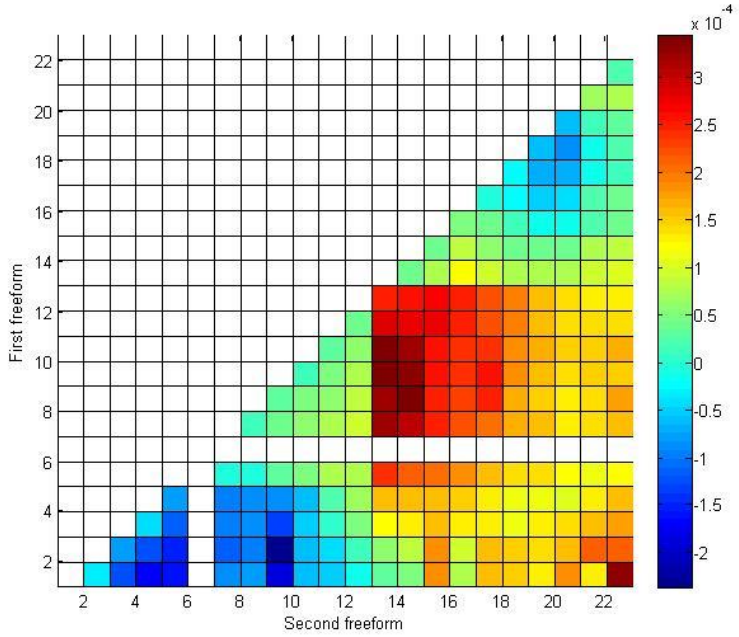



**Figure 35.** Influence of freeform sequence for the Scheimpflug system. Considering the symmetry of operation, the lower half of the diagram equals to the same part of Figure 34 minus its upper half ($M_{ij}^{1} = M_{ij} - M_{ji}$). The upper half of the diagram is zero after the subtraction.

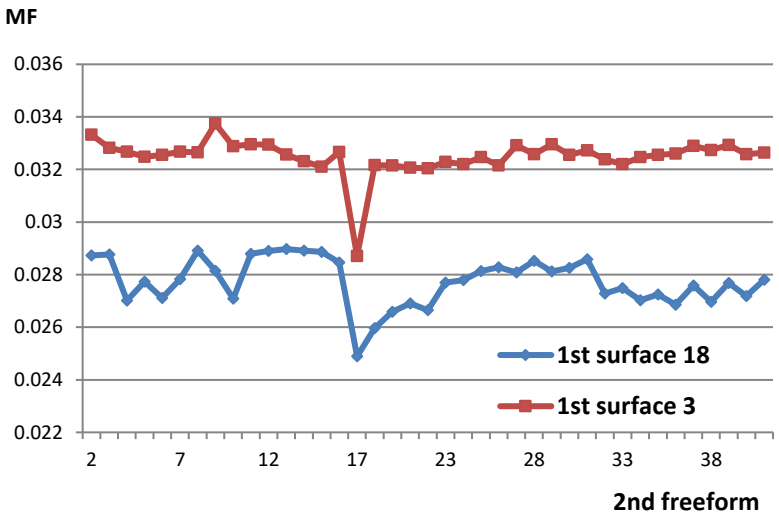



**Figure 36.** Lithography system, performance with a second freeform surface under different first freeform selections. The new best location for the next freeform surface is surface 17, when the first freeform surface is placed at surface 18 [64].

These concepts are applied to the lithography system. The two freeform mapping is shown in Figure 37, where we see again the best freeform positions are optimal positions of the one freeform analysis, surface 18 and 17, as indicated in Figure 21(a) and Figure 36. It happens that the best locations for the two freeforms are two mirrors close by. This seems to violate the observation we just proposed when optimizing the Scheimpflug system with two freeforms. However, a closer look shows that the two mirrors are next to the intermediate image planes, which are positioned between surface 17 and 18. The ray bundles from different fields are well separated on these mirror surfaces. The change of eccentricity is almost invisible due to the close vicinity of the mirrors. Moreover, these two mirrors break the system's rotational symmetry, which intuitively makes them good candidates for placing asymmetric freeforms.

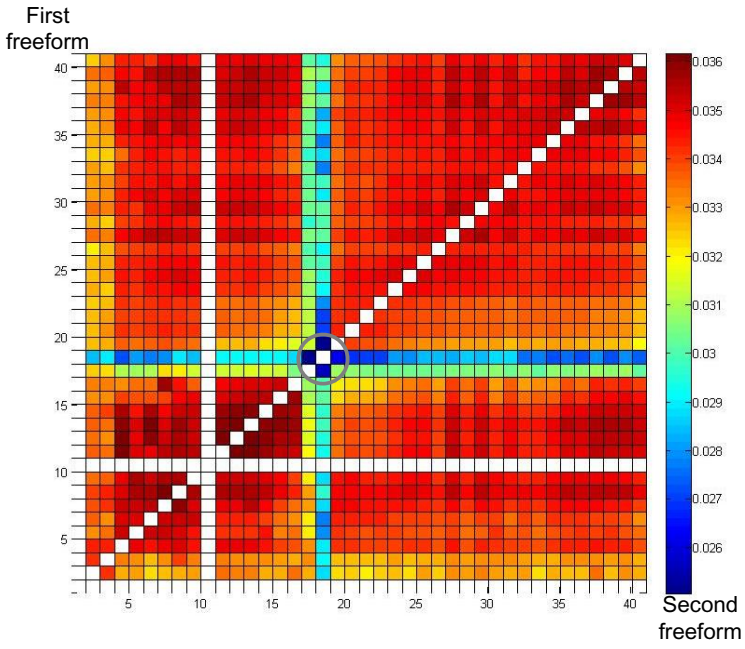



**Figure 37.** Lithography system, two freeform mapping. The y-axis corresponds to the first freeform surface index, and the x-axis corresponds to the second freeform surface index. The white elements in the diagram are invalid freeform combinations. Area of the optimal surface combinations is marked with a grey circle [64].

### 4.2.4 PROPOSED EFFICIENT WORK FLOW

To run all possible surface combinations for identifying the best freeform locations is truly very time-consuming. On an i7-3520M CPU@2.9 GHz, the run time for obtaining the two-freeform map of the lithography system shown in Figure 37 was more than 24 hours. A much simpler procedure that takes the single freeform result as a starting point for the final two freeform optimization, is proposed, which seems to be a quite general rule for the investigated systems.

As a general result, a workflow is proposed in Figure 38. This step by step workflow summarizes at best our experience to achieve a well-optimized system performance using freeforms. Multiple freeform surfaces are added one after another until the system specifications are reached or a better start system seems to be necessary. The analysis of the system's aberration contributions before each freeform implementation is important, which not only gives us information to determine the freeform order, but also helps predict the possible optimal freeform index. For example, a freeform surface close to the stop is needed if the system suffers from large field-independent aberrations; a freeform close to the field is helpful if large field-dependent aberrations exist. If both types of aberrations are large like for the Scheimpflug system, freeform is better placed in between.

The workflow described in Figure 38 suggests no structural changes made to the start system, therefore, parameters like lens thicknesses and lens spacing that have a lower impact on optimization results stay fixed. Once the optimal location of the freeform is determined, other parameters including all surface radii, surface conics, and potential asphere or freeform coefficients will be set as variables for achieving the best system performance. These extra variable are again stay fixed during the search of the optimal location for the next freeform.

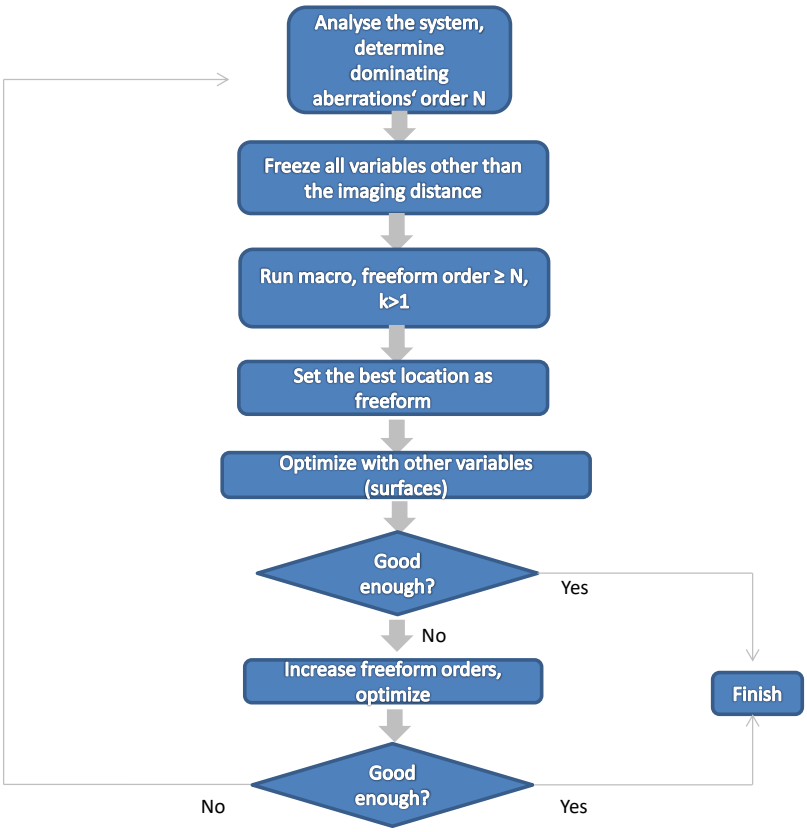



**Figure 38.** The final proposed workflow for implementing freeform surfaces in an optical system [64].

The preferred $\chi$ range for the selection of the freeform location in case of different aberration types is summarized in Table 4. For the correction of field independent component of the aberration, as well as for spherical aberration, the $\chi$ of the optimal freeform location is close to -1, i.e. the pupil location. For the correction of field dependent component of the aberration, as well as for the correction of distortion and telecentricity, the $\chi$ value of the optimal freeform location is close to 1. Both astigmatism and coma have field independent and dependent components, their correction depends on the constitution of both components.

**Table 4.** Preferred $\chi$ value for the selection of freeform location in case of different aberration types.

| | **Preferred $\chi$** |
|---|---|
| Field independent component (resolution) | Close to -1 |
| Field dependent component(resolution) | Close to 1 |
| Distortion | Close to 1 |
| Telecentricity | Close to 1 |
| Astigmatism | Not clear |
| Coma | Not clear |

| | |
|---|---|
| Spherical aberration | Close to -1 |

## 4.3 FREEFORM MANUFACTURABILITY

System tolerancing is a very important step of any optical system development, which is influenced by the sensitivity of each element. When considering freeform tolerancing, the positioning tolerances are similar to those of aspheric surfaces, with one extra consideration of the rotation of the element around the mechanical axis z. On the other hand, the tolerancing of the low-spatial-frequency figure error as illustrated in Figure 2 is much more complex and needs more experience. This aspect can only be discussed respectively for individual designs. Simple solutions to freeform tolerancing cannot be provided in this work, what can be done as an alternative is the first evaluation of freeforms' manufacturability.

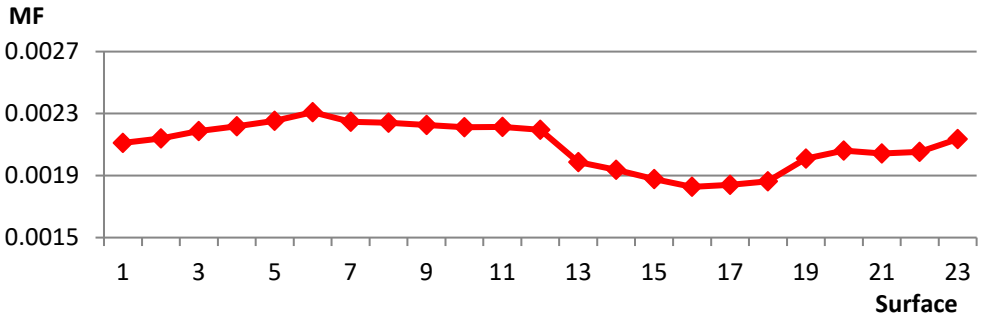
As mentioned in section 2.3.4, freeform manufacturability conditions determined by the machine are given for freeform sag, azimuthal gradient and azimuthal acceleration. These parameters decide whether a freeform can be manufactured under current technical constraints or not.

While the freeform sag defined in manufacturing is the complete surface sag with the rotational symmetric part subtracted, which excludes the rotational symmetric part of the freeform representation (e.g. defocus, spherical aberrations), I use in the context of this thesis the freeform sag definition of section 2.3.2. Merely the impact of the basic shape is excluded in our definition. With proper sampling, the PV value of the individual freeform sag is documented for further comparison.

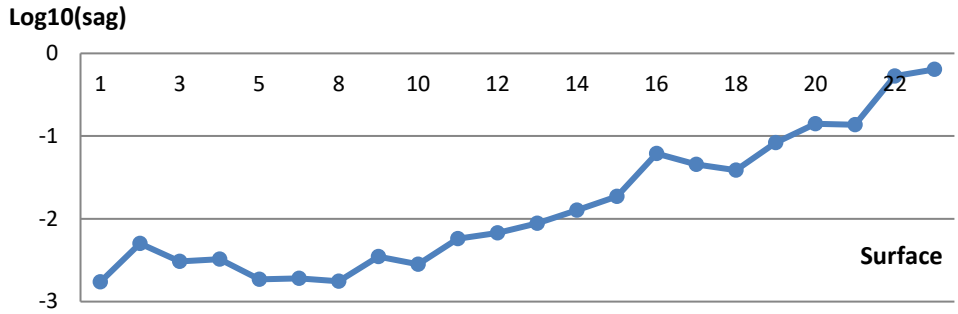
### 4.3.1 SAG CONTROL

Taking the Scheimpflug system as an example, the system's merit function values with varying freeform locations are given in Figure 39 (a), in which freeforms up to the $8^{th}$ order are placed at different surface locations for a fifty-cycle optimization without extra control of the freeform sag. The individual freeform sags are all

substantially smaller than the 6mm limit (in fact, below 0.5mm), whose distribution is indicated in Figure 39 (b).




(a)




(b)

**Figure 39.** Scheimpflug system, without sag control, system merit function values resulted from different freeform locations (a) and individual freeform surface sag values (b). The freeform sag is represented in logarithmic scale to in (b) to indicate the relative distribution.

Comparison between Figure 39 (a) and (b) shows that freeforms' sag values do not necessarily represent their aberration-correction contribution. Speculation can be made that the correction ability of the freeform is more location related, rather than shape related. Although the necessity for a sag control does not exist for the Scheimpflug system, I implement corresponding operands in the merit function to check their impacts on system performance. By integrating a self-written Zemax macro for freeform PV sag calculation in the merit function, the freeform sags can be controlled within each optimization cycle. The requirement for freeform sag is set to be not larger than 0.5mm, which has already been met by all freeforms.

As expected, extra constraints show no impacts on freeform sag values after optimization. Nevertheless, the merit function values undergo an equal percentage of decrease as represented in Figure 40. The differences between merit function values obtained with sag control and without sag control are divided by the merit function values without sag control in Figure 40. This parameter remains mostly constant for all surface locations, other than the last surface. The small spacing between the last surface and the image plane, and the requirement for distortion control with a tilted image can contribute to the abnormal behavior of the last surface, which is also seen from the large sag value of this surface in Figure 39 (b).

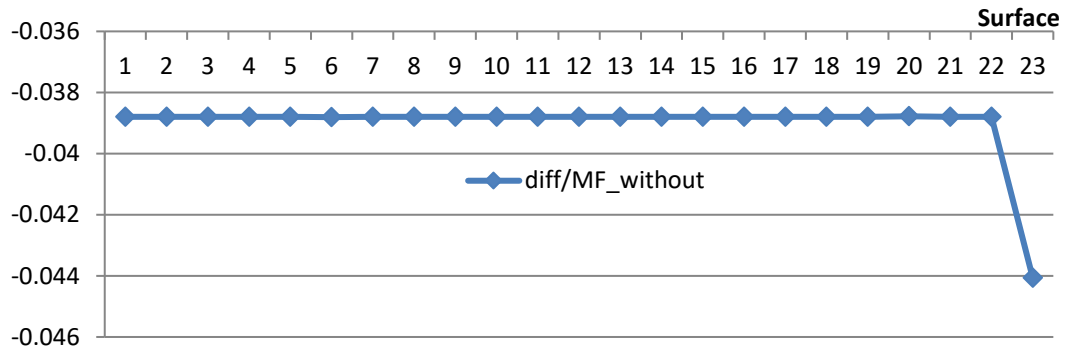



**Figure 40.** Scheimpflug system, relative change of the merit function value with sag control.

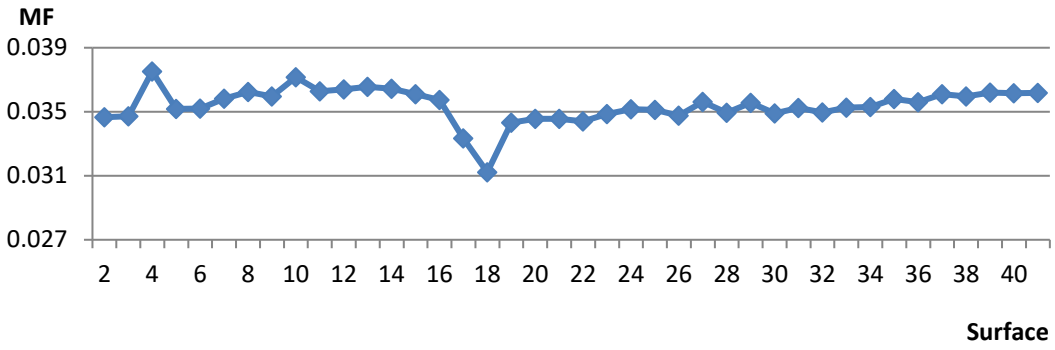
The same investigation is done for the lithography system. In Figure 41, merit function values with varying freeform locations are indicated, as well as the individual freeform sag values. For the lithography system, a similar finding is observed, where no obvious correlation can be established between Figure 41 (a) and (b).

With a proper sag control of smaller than 0.9mm, the new system performance represented by the new merit function values is indicated in Figure 42 (a), whereas the differences of the merit function values are plotted in Figure 42 (b). For surface locations, where large freeform sag values exist in the absence of proper sag control (e.g. surface 2, 10, 30, 31), the system performance becomes worse with decreased freeform sags. On the other hand, the merit function values decrease for nearly all surface locations, including those whose freeform sags fall within the range. This agrees with the finding concluded from the Scheimpflug example. A conclusion can

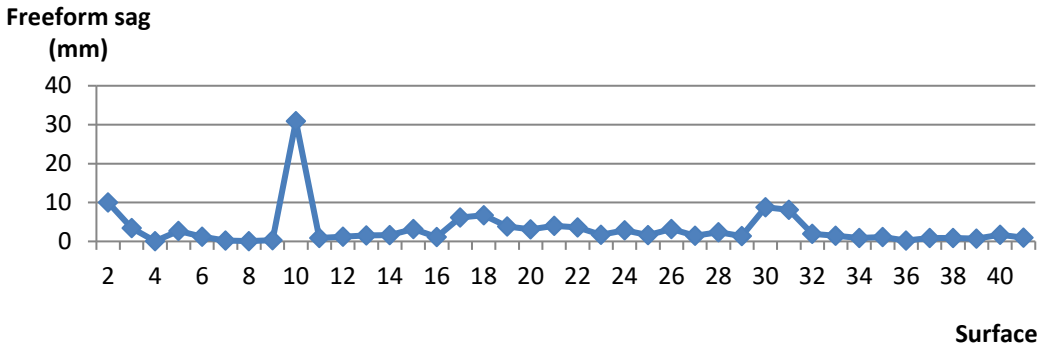
be drawn, that the value of the merit function is sensitive and related to the slight change of its constitution.

If a red line is drawn in Figure 42 (b) to estimate the baseline of the merit function value with unaffected system performance, only surface locations with originally large sags are above this line. Although the system performance is worse with controlled freeform sag values for these locations, the amount of increase of the merit function compared with the decrease of the PV sag value is trivial for this example.

Agreeing with the results obtained from section 4.2.2, the current optimization algorithm is shown to be suboptimal for freeform optimization, which tends to increase freeform coefficients unlimitedly to achieve better system performance. This sag increase in systems with minor remaining aberrations is not trivial, but becomes critical in more challenging systems, as indicated in the lithography system.

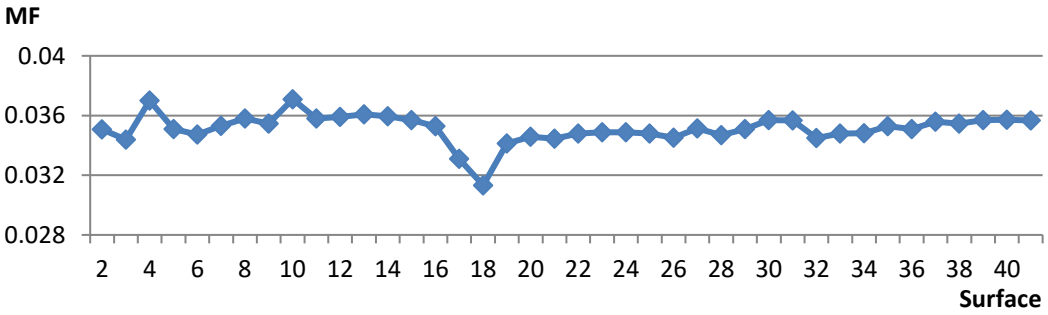



(a)

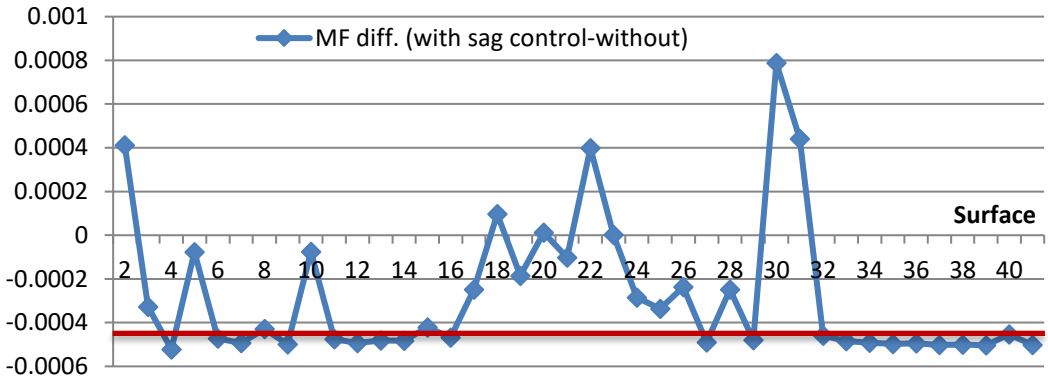



(b)

**Figure 41.** Lithography system, without sag control, system merit function values resulted from different freeform locations (a) and individual freeform surface sag values (b)




(a)




(b)

**Figure 42.** Lithography system, with sag control, system merit function values resulted from different freeform locations (a) and merit function value differences (merit function value with sag control minus merit function value without sag control) (b). The red line is drawn to estimate the base line of the merit function value with unaffected system performance.

Considering the long running time of the optimization process with freeform sag control, which is resulted from the long call-and-run time of the macro, the manufacturability control step can either be done with less sampling points, consequently less accuracy, or be done later. As indicated by the lithography system, the performance loss due to decreased surface sag is potentially trivial. Moreover, the merit function curve without sag control resembles that of the case with sag control as indicated in Figure 41 (a) and Figure 42 (a). This resemblance makes it sufficient to use the merit function curve without sag control for the freeform location assessment step.

### 4.3.2 GRADIENT AND ACCELERATION CONTROL WITH Q-POLYNOMIAL

Other than the restriction on freeform PV sag, which is given by the maximum vertical moving range of the tool, gradient and acceleration range of the freeform shape in the azimuthal direction are equally, if not more, critical. The Q polynomial, derived by Forbes [36, 37], indicated in Eq. (14), offers easy access to freeforms' RMS slope shown in Eq. (16). Starting from Eq. (15), the gradient and acceleration of the freeform departure part of the Q-polynomial representation in the azimuthal direction are derived in Eq. (54) and (55) respectively.

$$\frac{\partial \delta}{\partial \theta} = \sum_{m=1}^{M} \mu^m \sum_{n=0}^{N} \left[ -m a_n^m \sin(m\theta) + m b_n^m \cos(m\theta) \right] Q_n^m(\mu^2) \tag{54}$$

$$\frac{\partial^2 \delta}{\partial \theta^2} = \sum_{m=1}^{M} \mu^m \sum_{n=0}^{N} -m^2 \left[ a_n^m \cos(m\theta) + b_n^m \sin(m\theta) \right] Q_n^m(\mu^2) \tag{55}$$

The acceleration in azimuthal direction is similar in format to Eq. (15) with an extra factor $-m^2$. It can, therefore, be related directly to the freeform sag. The gradient in azimuthal direction is trickier, but since the mean square slope indicated in Eq. (16) is more readily accessible, the mean square slope can be used instead for manufacturability estimation. These values serve as first estimates for freeforms' manufacturability according to the constraints given in section 2.3.4.

These measures for testability and manufacturability are simple yes or no criteria, which give indications of whether the freeforms can be manufactured and tested or not. They do not reflect the freeform form error tolerance directly. The complete manufacturability of the freeforms can only be given with the analysis of the tool path based on the surfaces' freeform sag distribution. With the ongoing development of freeform manufacturing, the manufacturing characterization of freeforms can be better described in the future when more experience is gained.

Nevertheless, the freeform sag and root mean square slope of the freeforms are evaluated on the example Scheimpflug system as indications for manufacturability. The freeform performance with respect to its location can also be shown with a different freeform representation.

In Figure 43, the comparison of freeform performance between two different surface representations, Zernike fringe sag and Q-polynomial is given. As can be observed, no clear difference exists between these two representations, except for minor variation at the last surface. This finding is consistent with the published work [25]. Freeform PV sags of two different representations using the same optimization procedure are indicated in Figure 44. A large difference in sag clearly exists for the last system surface, whereas the sags of other surfaces are in comparable ranges. The potential issues of placing the freeform at the last surface have been addressed formerly in section 4.3.1. This aspect is once again shown here with another example.

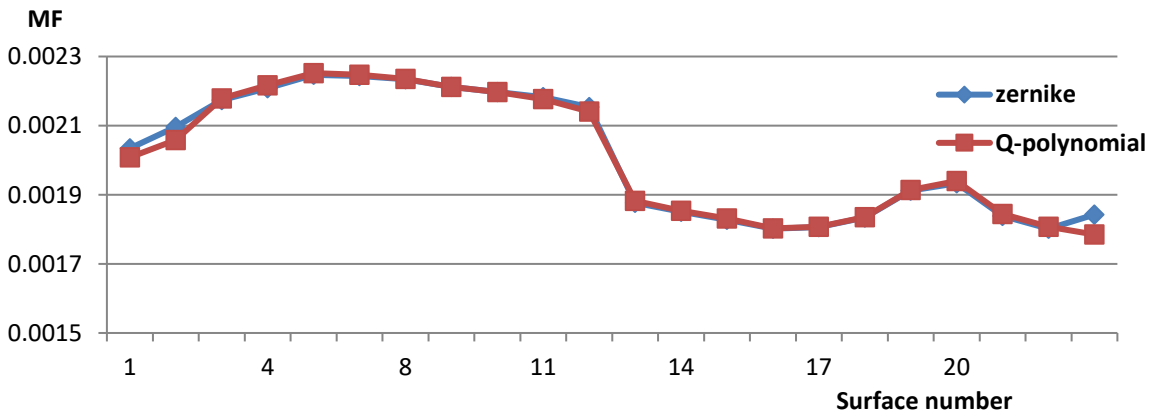


**Figure 43.** Scheimpflug system, freeform performance comparison between two different surface representations, Zernike fringe sag and Q-polynomial. The x-axis is the freeform location, the y-axis is the merit function value.

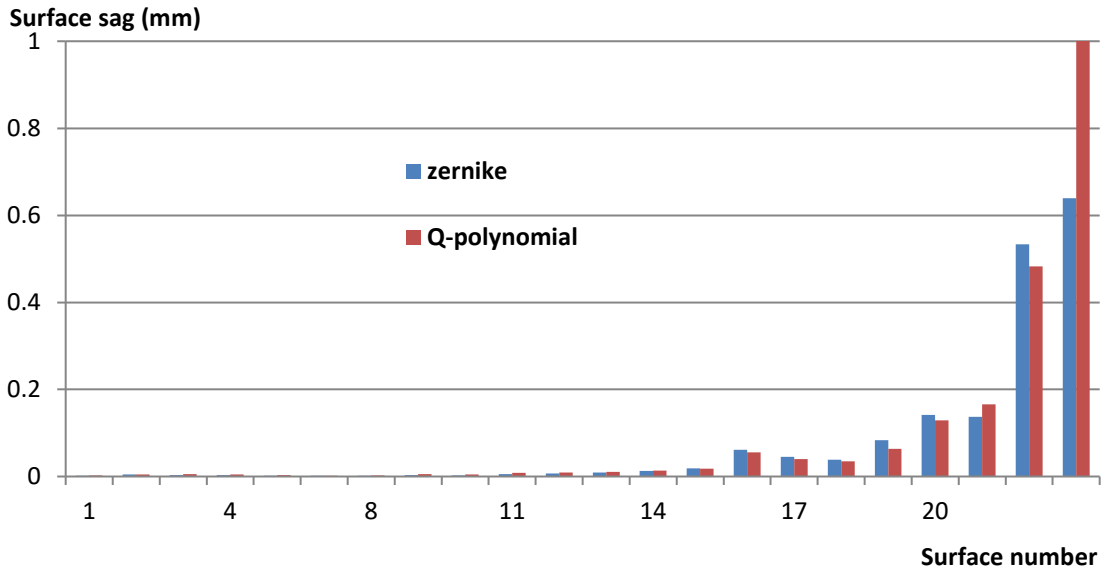



**Figure 44.** Scheimpflug system, freeform sag comparison between Zernike fringe sag and Q-polynomial. The x-axis is the freeform location. The y-axis is the freeform PV sag, whose range is readjusted to accommodate surfaces in the front. The freeform sag of the last surface is 4.7mm for the Q-polynomial representation.

Detailed information of the freeform RMS slopes and PV sags of different surface locations is given in Table 5. The mean square slope (MS_slope) corresponds to Eq. (16), which is calculated directly from freeform coefficients. Since the magnitude of the coefficient is closely related to the magnitude of the normalization radius, this information is also given in the table. In principle, for the same freeform, the larger the normalization radius, the larger the coefficients, and consequently the larger the MS-slope. What is also interesting for the Q-polynomial representation, is the RMS slope value. Based on the values of the RMS slope of this example, the general tendency seems to be, when the sag values are large, the slope values are also large even with smaller normalization radii. The last two surfaces with large sag and slope values are repeatedly shown to be less preferred from the manufacturability point of view.

Despite the simplified approaches used here for the freeform manufacturability assessment, the close relations between freeform PV sags and manufacturability constraints provide the possibility for a fast first estimation.

**Table 5.** Scheimpflug system, freeform manufacturability assessment with Q-polynomial. Freeform RMS slopes and PV sags, which are used to represent freeform topology, are given in the last two columns.

| Surface | MF | Resolution(μm) | Norm. radius (mm) | MS_slope | RMS_slope | PV_sag (mm) |
|---|---|---|---|---|---|---|
| **1** | 0.002008 | 0.002065 | 11.68794 | 0.000231 | 0.015198 | 0.00278 |
| **2** | 0.002058 | 0.002033 | 11.23529 | 0.001065 | 0.032628 | 0.004816 |
| **3** | 0.002179 | 0.001721 | 11.46791 | 0.002131 | 0.046163 | 0.005767 |
| **4** | 0.002217 | 0.001538 | 11.6406 | 0.001944 | 0.044092 | 0.004989 |
| **5** | 0.002252 | 0.001324 | 13.81294 | 0.000863 | 0.02937 | 0.00327 |
| **7** | 0.002247 | 0.001337 | 16.92461 | 0.000146 | 0.012073 | 0.001674 |
| **8** | 0.002235 | 0.001413 | 17.17781 | 0.000526 | 0.022945 | 0.002817 |
| **9** | 0.002212 | 0.001513 | 16.93647 | 0.001998 | 0.044694 | 0.005559 |
| **10** | 0.002197 | 0.001592 | 16.76642 | 0.001397 | 0.037375 | 0.004831 |
| **11** | 0.002177 | 0.001667 | 13.89925 | 0.005224 | 0.072274 | 0.008597 |
| **12** | 0.002141 | 0.001805 | 11.06115 | 0.005871 | 0.076624 | 0.009198 |
| **13** | 0.001882 | 0.002391 | 10.21708 | 0.005333 | 0.073026 | 0.011131 |
| **14** | 0.001854 | 0.002444 | 10.8223 | 0.006175 | 0.078579 | 0.014199 |
| **15** | 0.001831 | 0.002537 | 11.02953 | 0.005473 | 0.073983 | 0.018882 |
| **16** | 0.001803 | 0.00268 | 10.43058 | 0.037905 | 0.194692 | 0.057517 |
| **17** | 0.001807 | 0.002673 | 8.876731 | 0.035543 | 0.188529 | 0.041569 |
| **18** | 0.001835 | 0.002653 | 8.644613 | 0.031165 | 0.176536 | 0.035309 |
| **19** | 0.001914 | 0.002382 | 7.705161 | 0.078246 | 0.279724 | 0.070414 |
| **20** | 0.00194 | 0.002214 | 7.098853 | 0.280352 | 0.529483 | 0.130174 |
| **21** | 0.001844 | 0.002165 | 5.700223 | 0.129921 | 0.360445 | 0.168193 |
| **22** | 0.001807 | 0.00182 | 5.395272 | 1.689115 | 1.29966 | 0.488055 |
| **23** | 0.001785 | 0.001798 | 3.619294 | 72.82138 | 8.533544 | 4.732469 |

# 5 Evaluation of the methods with further test systems

In the previous chapter, the optimization strategies involved in freeform implementation are discussed mainly based on two example systems. The findings, as well as the suggested workflow, are obtained from a small number of systems. Though it is apparently unlikely to cover all system types in the context of this work, we would like to discuss the results and evaluate the findings and proposed procedures obtained so far with more examples. Moreover, system performance improvement with freeforms implemented according to the workflow will also be demonstrated.

## 5.1 ALVAREZ LENS SYSTEM WITH NO SYMMETRY

Applying the macro indicated in Figure 18 on the Alvarez system shown in Figure 24, we obtain the system performance curve indicated in Figure 45. The merit function values of four different freeform locations are given in this figure. It is noticed that the system has better performance with freeforms placed at the first Alvarez lens (surface 3, 4). The explanation for this is simple. The residual dominating aberrations of this Alvarez lens system are induced aberrations caused by the finite distance between two plates. The magnitude of the higher order induced aberration relates closely to the low order intrinsic aberrations. It is therefore beneficial to correct aberrations early in the system to reduce the magnitude of induced aberrations. The optimal location of the first freeform is surface 4 according to Figure 45.

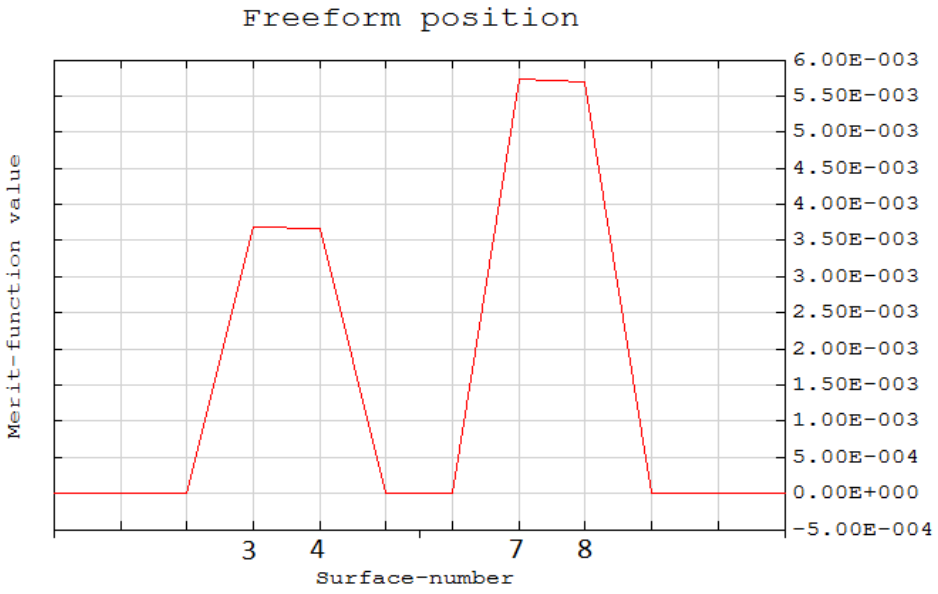



**Figure 45.** Alvarez lens, the system performance represented by merit function values is given for different freeform positions.

When it comes to the selection of the location of the second freeform, the optimal solution varies as the magnitude of the system's residual induced aberration changes. As indicated in Table 6, the two Alvarez plates are placed with different spacing d varying from 2mm to 0.4mm. As the spacing decreases, the residual induced aberration of the original system decreases. The optimal second-freeform location switches from surface 8, which is further away from the first freeform, to surface 7. This switch happens roughly at d=0.67mm according to Figure 46. This figure also shows that there is no contradiction between the conclusion drawn in the last chapter regarding the placement of the second freeform and the induced aberration nature of the Alvarez lens. When two Alvarez lenses are placed far away from each other, the induced aberrations caused by the two $6^{th}$ order freeforms of the original system dominate. The second freeform needs to be placed away from the first freeform, to have the maximum correction ability, as also indicated in Figure 32 for the Scheimpflug system. However, when the residual induced aberrations originated from the original setup become small as two Alvarez lenses approaching each other, the large spacing between two manually implemented freeforms becomes an issue. The newly introduced induced aberrations due to the larger spacing outweigh the benefits, it is therefore preferred to place the second freeform closer to the first

freeform. This transition from dominating old induced aberrations to dominating new induced aberrations is shown clearly in Figure 46.

**Table 6.** Alvarez lens system, freeform performance with varying lens spacing. The second manually implemented freeform is placed at all residual locations, surface 3, surface 7 and surface 8. The respective system performance represented by the merit function value is given for each location. The differences between placing the freeform at surface 7 and surface 8 are compared for different lens spacing.

| | d=2mm | d=1.5mm | d=0.8mm | d=0.7mm | d=0.6mm | d=0.5mm | d=0.4mm |
|---|---|---|---|---|---|---|---|
| s3 | 0.01000924 | 0.00679901 | 0.00288217 | 0.00246033 | 0.00212969 | 0.00193411 | 0.00191252 |
| s7 | 0.00998317 | 0.00678122 | 0.00288111 | 0.00246029 | 0.00212900 | 0.00193033 | 0.00190311 |
| s8 | 0.00997844 | 0.00677789 | 0.00288050 | 0.00246016 | 0.00212932 | 0.00193096 | 0.00190378 |
| s7-s8 | +4.73E-06 | +3.33E-06 | +6.1E-07 | +1.3E-07 | -3.2E-07 | -6.3E-07 | -6.7E-07 |

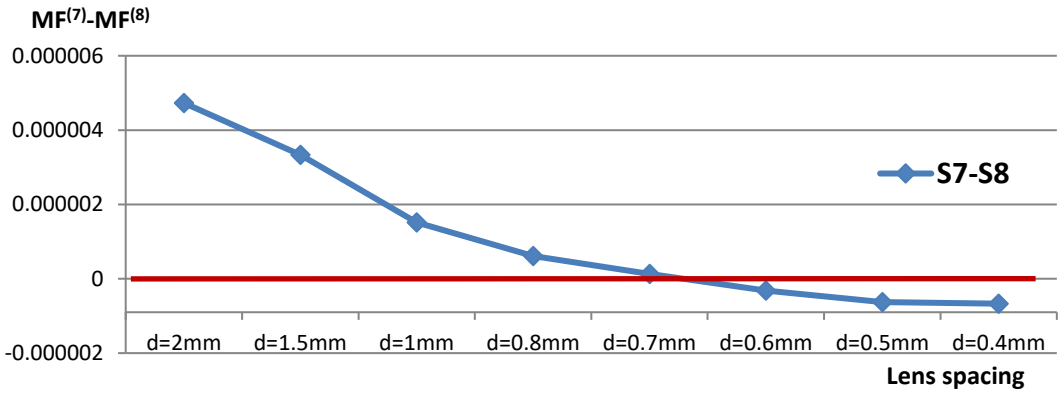



**Figure 46.** Alvarez lens, merit function value versus varying lens spacing. This figure visualizes the last row of Table 6, and shows the switch of the optimal freeform location from surface 8 to surface 7 as two Alvarez lenses approaching each other.

## 5.2 SCHEIMPFLUG SYSTEM WITH PLANE SYMMETRY

So far, the Scheimpflug system in chapter 4 has been investigated using only a limited number of system variables. The proposed workflow is now validated in Figure 38 with the optimization of the Scheimpflug system using all surface radii of curvature. The optimization includes freeform placement at surface 16 and surface 1 with a successive strategy as suggested by the workflow. Figure 47 shows the final results of our optimization, where we start with a non-diffraction-limited system while finalizing the design process as shown in Figure 47(b) with a nearly

diffraction-limited performance. The whole optimization process has been finished within ten minutes.

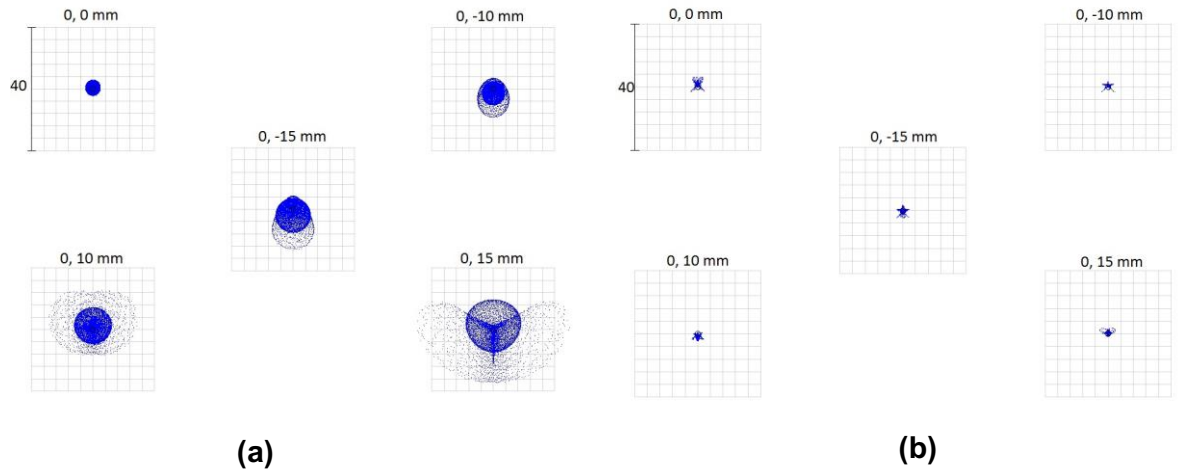



**Figure 47.** Scheimpflug system spot diagram of (a) the start system without freeforms and (b) the optimized system with two freeforms. Both diagrams have the same reference box size. The Airy radius equals to 0.81 μm for (b) [64].

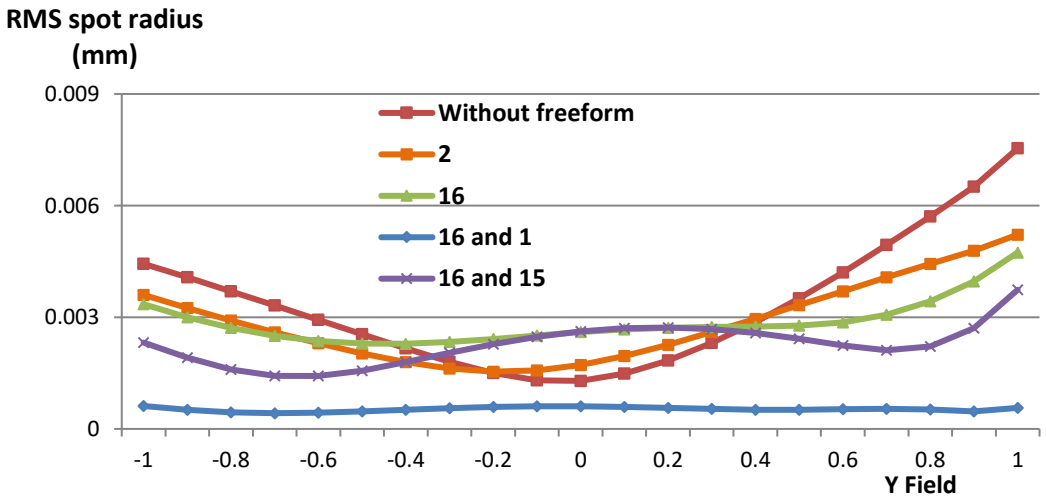



**Figure 48.** Scheimpflug system, RMS spot radius in relation to selected freeform locations. The red curve is the field performance of the system along y with only spherical surfaces. With the best two-freeform selection surface 16 and surface 1 chosen, the overall spot size is not only decreased for the whole field, but also has an improved homogeneity [64].

Figure 48 indicates the improvement in the RMS spot radius for different configurations without freeform, with one freeform and with two freeforms. As expected, freeform one and two placed at surface 16 and surface 1 perform better than other combinations. In addition, we observed an improved homogeneity and

an improved RMS spot radius over the complete field. This clearly indicates that the proposed workflow leads to a fast selection of a first guess of the freeforms. The results are summarized in Table 7, where we see substantial improvement from the original system. The $\chi$ of the first freeform 16 is -0.4, which is closer to the field plane than the second freeform 1 ($\chi$=-0.7). The field uniformity is therefore improved better by a factor of 2.8 with the first freeform, while the averaged resolution is improved by a factor of 5.4 with the second freeform.

**Table 7.** Scheimpflug system resolution improvement with freeforms.

| | **Original** | **1 freeform** | **2 freeform** |
|---|---|---|---|
| Averaged RMS radius $\phi_{ave}$ | 3.33 μm | 2.86 μm | 0.53 μm |
| Non-uniformity $\phi_{\max} / \phi_{\min}$ | 5.84 | 2.07 | 1.48 |

## 5.3 LITHOGRAPHY SYSTEM WITH PLANE SYMMETRY

The example lithography system shown in Figure 15 is altered from the original patent using only spheres. In reality, as shown in previous sections, the system suffers from large residual aberrations due to the elimination of the use of aspheres. This drawback can be overcome either by introducing multiple freeforms with potentially large sags, or by using aspheres for pre-correction. The system after re-optimization has two aspheres up to the 16$^{th}$ order, and six aspheres up to the 14$^{th}$ order. Despite the performance enhancement, room for improvement still exists. Locations of the aspheres are marked red in Figure 49. With freeform locations selected according to the workflow in Figure 38, the layout of the optimized system is given in Figure 50. The optimized system has two freeforms placed at surface 18 and surface 25, with one surface closer to the field and one surface closer to the pupil. System performance of both systems is listed in Table 8, where improvement in all aspects up to a factor of 10.5 is observed for the freeform system. The decoupling of the tangential and sagittal correction of the freeforms helps tremendously with the correction of astigmatism and coma, as shown in Figure 52, where the improvement of these two is observed over the complete y field.

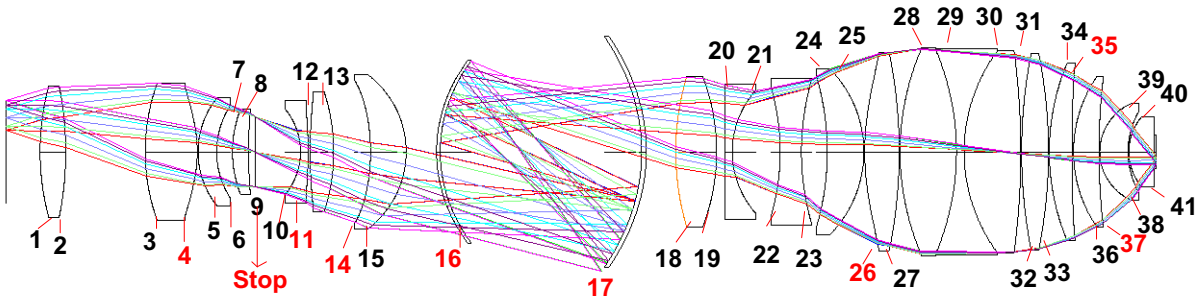



**Figure 49.** Layout of the lithography system with eight aspheres. Asphere locations are marked red.

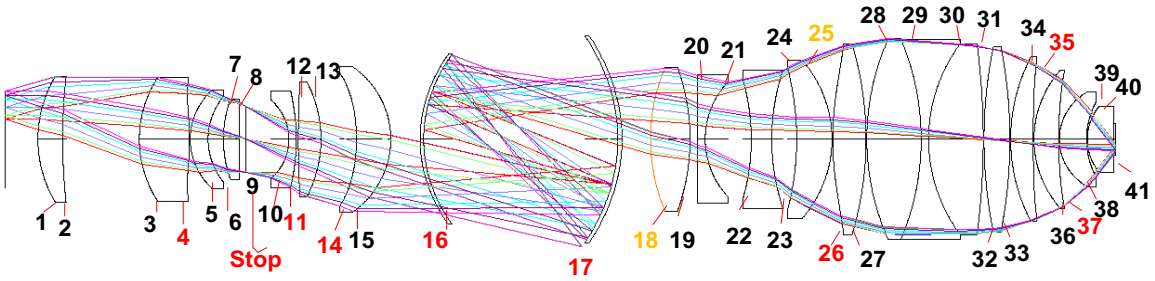



**Figure 50.** Layout of the optimized lithography system with eight aspheres and two freeforms. Asphere locations are marked red, and freeform locations are marked yellow.

**Table 8.** Lithography system, performance comparison between the original system and the optimized system. The original system has eight aspheres, and the optimized system has two additional freeforms placed at surface 18 and surface 25. Improvement factor up to 10.5 is observed for the freeform system.

| | Distortion (μm) | Telecentricity (°) | Averaged RMS spot radius(μm) |
|---|---|---|---|
| **Original system** | 0.038 | 0.386 | 0.446 |
| **Optimized system** | 0.0036 | 0.205 | 0.047 |
| **Improv. factor** | 10.5 | 1.9 | 9.5 |

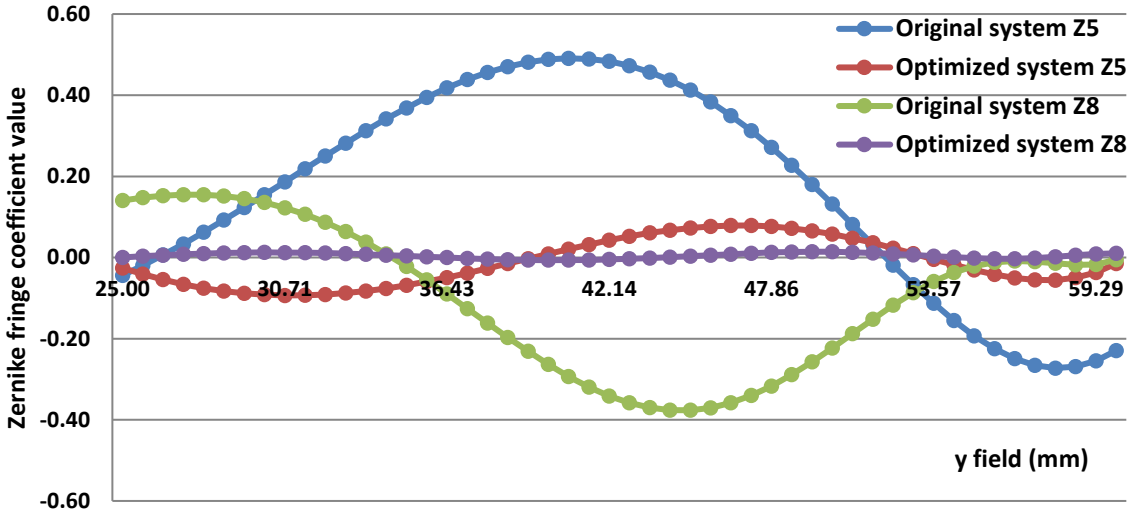



**Figure 51.** Z5 and Z8 distribution over the y field for the original lithography system and the optimized lithography system. Both aberrations are improved tremendously for the optimized system.

## 5.4 SPECTROMETER SYSTEMS

Spectrometer systems with broken system symmetries are good examples for freeform implementation, especially in the case of mirror systems, where decenters and tilts of individual mirror elements exist to avoid central obscuration. In this subsection, several spectrometer systems will be investigated under the same specifications indicated in Table 9 [66, 67]. The near-infrared hyper-spectral spectrometer systems have an f/3, and are designed with maximum throughput for earth environmental observation. The detector pixel pitch is selected to be 24 μm for the examples, whose size is used as a reference for distortion evaluation. The impacts of freeform location selection for spectrometer systems will be analyzed in the context of this section.

**Table 9.** Spectrometer specifications.

| | **VNIR** |
|---|---|
| Wavelength (nm) | 420 – 1000 |
| F number | 3 |
| X field in height (mm) | ±12 |
| Spectral resolution (nm) | 6.5 |
| Smile / keystone distortion | <20% of a pixel pitch |
| Pixel geometry (μm) | 24 x 24 |

| | |
|---|---|
| MTF | >25% at 20.8 cyc/mm across track<br>>16% at 20.8 cyc/mm along track |
| Grating efficiency | ≥70% |

### 5.4.1 DOUBLE TMA SYSTEM

The first spectrometer system is a reflective Double TMA system with six mirrors and one plane grating from [67]. The original system is composed of two TMA systems employing one reflective plane grating separating the spectral information. Lacking the guidance of the workflow, the locations of non-spherical surfaces were selected randomly in the process of system evolution. With the stop located at mirror 2 (M2), the first TMA system (M1 to M3) uses three 4th order aspheres, and the second TMA (M4 to M6) uses three freeforms up to the 10th order.

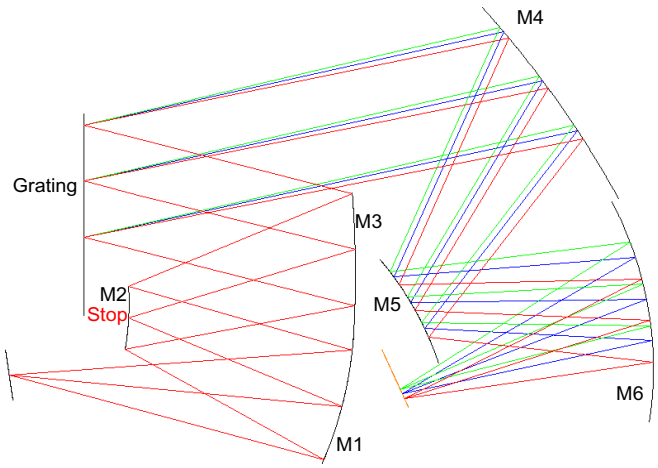



**Figure 52.** Layout of the double TMA system with a reflective grating [69].

Findings from the last chapter are applied on a modified version of the double TMA system containing only conic surfaces. With the assistance of the workflow, the freeform positions are chosen efficiently only at M6 and M4 up to the 10th order [69]. The new system consists of two freeforms and four conic surfaces. With the decreased number of both freeforms and aspheres, the system is still able to maintain comparable resolution and distortions. The comparison of system performance between the original system and the new system is given in Table 10. This example demonstrates that the workflow concluded from multi-lens systems works also for the multi-mirror system.

**Table 10.** Double TMA system, comparison between the optimized system and the original system. The distortion types are measured with respect to one pixel pitch 24 μm. The original system uses three aspheres for the first TMA system, and three freeforms for the second TMA system. The new system is able to maintain a comparable set of specifications with four conic surfaces and two freeforms placed at the second TMA.

| | F number | Freeform | Asphere | Conic surface | Keystone distortion | Smile distortion | Averaged RMS spot radius |
|---|---|---|---|---|---|---|---|
| **Original system** | 3 | 3 | 3 | 0 | <10% | <7% | 2.4 μm |
| **New system** | 3 | 2 | 0 | 4 | <9% | <5% | 2.3μm |

### 5.4.2 DOUBLE TMA SYSTEM WITH A TRANSMISSIVE GRATING

Another spectrometer system example is a double TMA system similar to Figure 52, but with a transmissive plane grating. With the stop placed at the grating, the system has a quasi-symmetric layout that is beneficial for distortion correction. The original system in Figure 53 has three freeforms (at the second TMA system) and three conic surfaces [67]. With the elimination of the use of freeforms, the system performance drops drastically as indicated in Table 11. The system without freeforms suffers mainly from the field independent astigmatism, and the optimal freeform location is, according to the workflow, at M3, which is close to the stop. The new system with only one freeform has a resolution improvement factor of more than two, and a distortion improvement factor of three. Although the new system still has worse resolution compared with the original system, the comparable or even better performance in distortion, along with the decreased freeform number, makes the new system a good application of the proposed workflow.

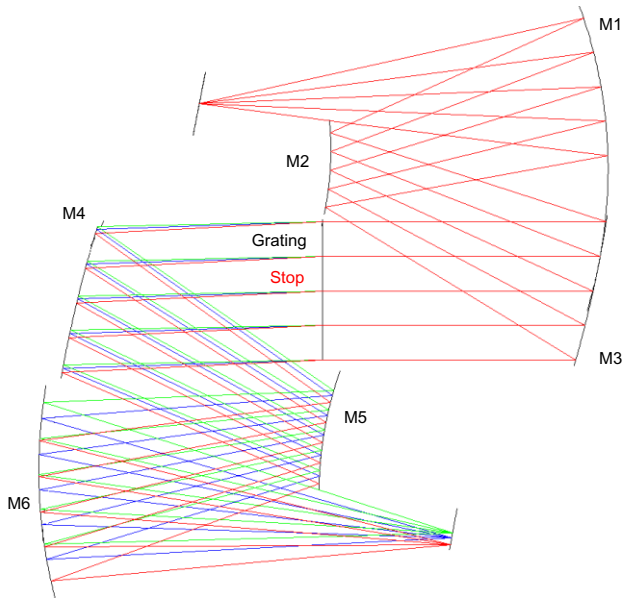



**Figure 53.** Layout of the double TMA system with a transmissive grating.

**Table 11.** Double TMA system with a transmissive plane grating, comparison between the original system, system without freeforms and the new system with less freeforms. The distortions are measured with respect to one pixel pitch 24 µm. The new system utilizes less freeforms, but exhibits better distortion correction.

| | f/# | Freeform | Asphere | Conic surface | Keystone distortion | Smile distortion | Averaged RMS spot radius |
|---|---|---|---|---|---|---|---|
| **Original system** | 3 | 3 | 0 | 3 | <2.5% | <10% | 1.6 µm |
| **Without freeforms** | 3 | 0 | 0 | 6 | <7% | <12% | 7.4µm |
| **New system** | 3 | 1 | 0 | 5 | <2.5% | <4% | 3.8µm |

### 5.4.3 DYSON SYSTEM

The catadioptric Dyson system shown in Figure 54 is different in composition from the above-shown spectrometers, having no collimator. The varying incidence angles on the grating due to the absence of the collimator leads to additional degradation of the spectral resolution and efficiency [48]. Moreover, the curved grating substrate contributes significantly to the system's aberration, which makes the consideration

of the finite substrate thickness no longer negligible. Based on previous work [67], a realistic 6.35mm thick quartz substrate is considered for the system in Figure 55.

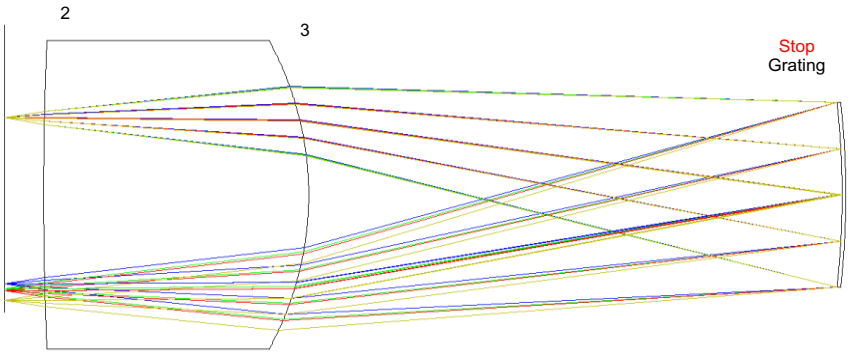



**Figure 54.** Layout of the Dyson system without substrate.

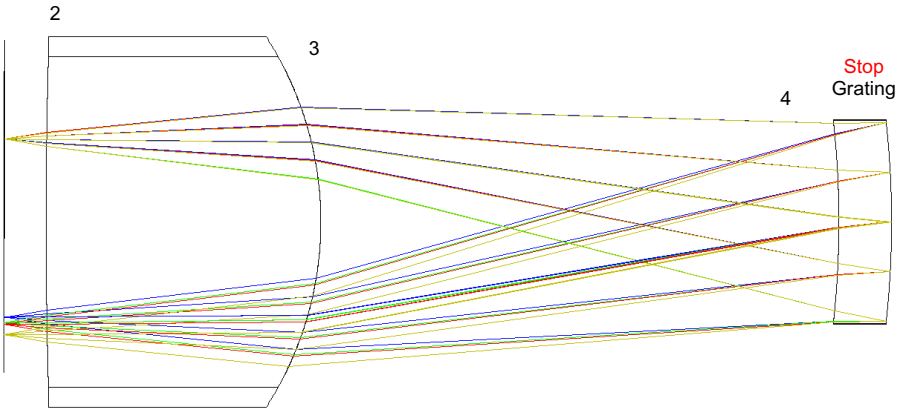



**Figure 55.** Layout of the Dyson system with a substrate.

Both Dyson systems share the same specifications as other spectrometer systems. The version with the quartz substrate has larger spot size (i.e. worse resolution) as indicated in Table 13. With freeforms placed at different locations, surface 2 and surface 3, the resulting system performance represented by distortions and resolution is shown in Table 12 for the Dyson system without substrate, and in Table 13 for the Dyson system with a substrate. All surfaces are used in the optimization process to achieve the best results.

In both cases, system resolution is improved when freeforms are placed at both locations. Nevertheless, the distortions are slightly increased, though still in specification, when the freeforms are placed at surface 3. When the freeforms are

placed at surface 2, due to the small spacing between this surface and the image plane, similar effects are observed as for the Scheimpflug system, where larger freeform sags occur. Moreover, system distortion experiences a larger increase in this case, due to the larger violation of the system symmetry. With appropriate performance criteria chosen, which in this case is the consideration of both distortion and resolution, the better location for freeform implementation is surface 3 for both systems.

The information of the freeform sags is provided for this system type to depict potential issues when freeforms are placed very close to the image plane. As also stated in previous sections, without specific restrictions stated in the merit function, the current optimization algorithm is not able to optimize freeform sags efficiently. This drawback is especially noteworthy when freeforms are placed near the image.

**Table 12.** Dyson system without substrate

| | f/# | Freeform | Keystone distortion | Smile distortion | Averaged rms spot radius | Freeform sag |
|---|---|---|---|---|---|---|
| **Orignal system** | 3 | 0 | <6% | <6% | 8.7μm | 0 |
| **freeform@2** | 3 | 1 | <18% | <7% | 2.8μm | 0.46mm |
| **freeform@3** | 3 | 1 | <10% | <7% | 3.8μm | 0.13mm |

**Table 13.** Dyson system with substrate

| | F number | Freeform | Keystone distortion | Smile distortion | Averaged rms spot radius | Freeform sag |
|---|---|---|---|---|---|---|
| **Orignal system** | 3 | 0 | <14.9% | <6% | 12μm | 0 |
| **freeform@2** | 3 | 1 | <25% | <10% | 3.4μm | 1.44mm |
| **freeform@3** | 3 | 1 | <14% | <10% | 4.8μm | 0.25mm |

### 5.4.4 DOUBLE PASS TMA SYSTEM

The double pass configuration is the configuration where the same sub-system is used twice under reversed ray directions. The grating is placed perpendicular to the ray bundles, providing the optimal condition for achieving the highest possible grating efficiency. One potential problem of the double pass configuration is the close distance between the object plane and the image plane. This problem can be overcome by tilting the grating slightly [67].

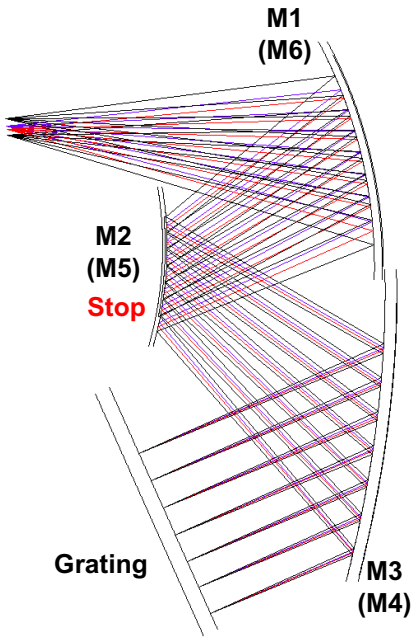



**Figure 56.** Layout of the double pass TMA system.

The double pass TMA system shown in Figure 56 has the system stop at M2. The asymmetric layout of this system is not ideal and leads to large distortions. The original system has three conic surfaces and large distortions and spot radius as presented in the second row of Table 14. Due to the bad choice of the start system, this example needs two freeforms to reach system specifications. All three two-freeform combinations and their resulting system performance obtained optimizing all surfaces are presented in Table 14. Clear differences between different freeform location selections are seen, where the M1-M2 combination gives the best overall system performance in both distortions and resolution. For example, the keystone distortion given by the best freeform combination is only 37.5% of that of the worst

freeform combination. The same is true for the resolution, where a much smaller RMS spot radius is possible with a better freeform location selection.

Since the system has only three surfaces with the stop placed in the middle, it is not clear to see the relation between the optimal freeform location and the pupil position. However, due to the large residual field-dependent and field-independent aberrations of the original system, two freeforms with one close to the stop and one close to the field are needed. As can been observed from Table 14, three optimized freeform systems all have much-improved performance (max. improv. factor 7.6) compared to the original system. The advantages of freeforms in correcting residual aberrations are obvious, and the selection of freeform locations is proven to be necessary with this example.

**Table 14.** Double pass TMA system, system performance comparison between the original system and systems with freeforms placed at different surface locations.

| | f/# | Freeform | Conic | Keystone distortion | Smile distortion | Averaged RMS spot radius |
|---|---|---|---|---|---|---|
| **Original system** | 3 | 0 | 3 | <45% | <50% | 33.3μm |
| **M1M3** | 3 | 2 | 1 | <15% | <18% | 6.4μm |
| **M1M2** | 3 | 2 | 1 | <6% | <14% | 4.4μm |
| **M2M3** | 3 | 2 | 1 | <16% | <16% | 6.6μm |

## 5.5 TELECENTRIC TWO-MIRROR SCAN SYSTEM

A telecentric x-y scan system is a good example of freeform implementation. Due to the difference in pupil positions where the scan mirrors are placed, the system has slightly different focal lengths in x and in y. It is, in fact, a weak anamorphic system with double plane symmetry. The original rotational symmetry does not exist anymore due to the introduced pupil separation.

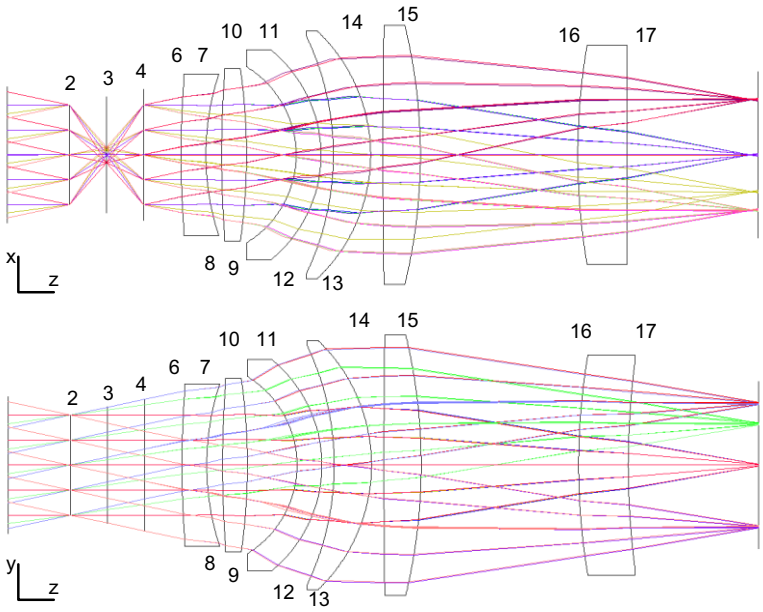



**Figure 57.** Layout of the original scan system with one biconic surface at surface 17.

Since the setting of separated pupil locations is not possible in Zemax, two ideal cylindrical lenses are implemented in x at surface 2 and 4 to translate the pupil position from surface 2 to surface 4 for the x-z plane. The layouts of the system in x-z and y-z plane are given in Figure 57. The pupil locates at surface 2 for the y-z plane, and locates at surface 4 for the x-z plane. The system has an f/3, and $\pm 12^\circ$ fields in both x and y. It works with e light (546.074 nm) and has one biconic surface at surface 17 (no conics, only different radii of curvature). The system is optimized to be telecentric in the image side, which means the exit pupil of the system should locate at infinity. To assure a reasonable spacing between the image plane and the last surface, the free working distance is set to be not smaller than 10mm for a focal length of 24mm in y and a focal length of 19mm in x.

The original system has a maximum grid distortion of -3.14%, an averaged RMS spot radius of 67.09 μm and a maximum violation of telecentricity of 0.406°. The system's largest aberration contribution is the primary astigmatism. With the analysis of the full field distribution of the systems' astigmatism (Zernike fringe coefficient Z5 and Z6) in Figure 58, we notice that the dominating aberration is the field independent astigmatism. According to nodal aberration theory, as well as our investigations discussed above, the freeform is better placed at locations close to the stop to correct

this type of aberration. In Figure 59, the system performance with different freeform locations (a), as well as surface eccentricity distribution (b) is given. The similarities between these two curves are observed, while the best freeform location is surface 6, which is the closest to the pupil position.

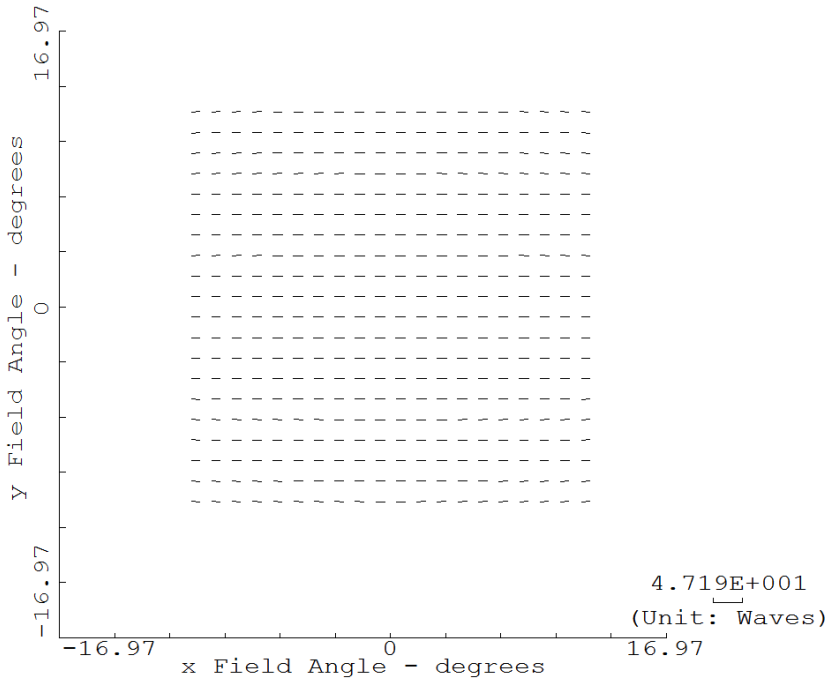



**Figure 58**. Full field distribution of astigmatism of the x-y scan system. The dominating aberration contribution is field independent astigmatism.

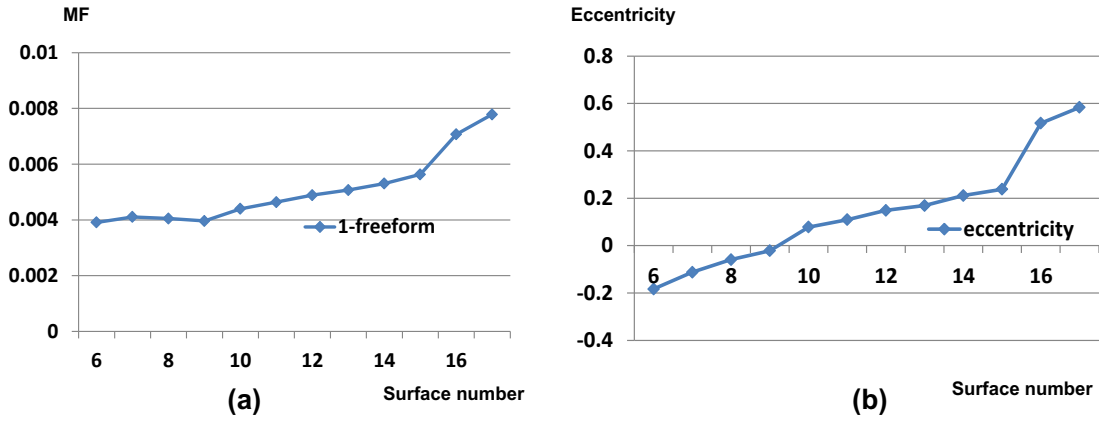



**Figure 59.** Scan system, selection of the first freeform location. System performance with respect to different surface locations is given in (a), and the eccentricity curve is given in (b).

After implementation of the first freeform placed at surface 6, the dominating residual aberration is the field dependent astigmatism according to the full field aberration distribution in Figure 60. The optimal location of the second freeform is therefore close to the field plane. As verified by Figure 61, the best locations to implement the second freeform are surface 16 or surface 17.

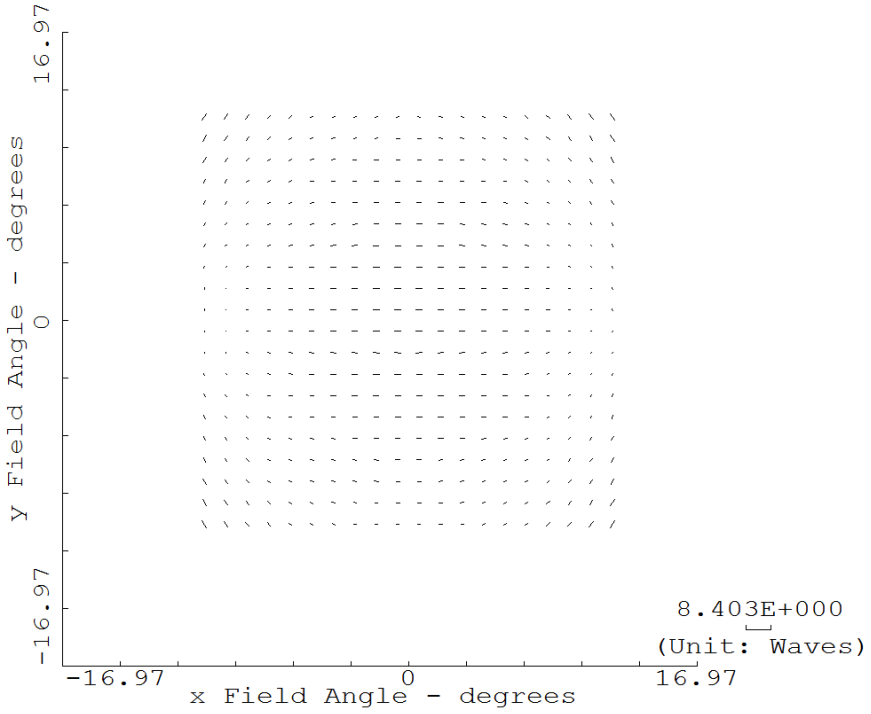



**Figure 60.** Full field distribution of astigmatism after implementation of the first freeform. The dominating aberration is field dependent astigmatism.

According to the results suggested by the workflow, surface 6 and surface 17 are replaced with freeform surfaces and the complete system is optimized using the same merit function. The performance of the optimized freeform system is shown in Table 15, and improvements in all aspects are indicated.

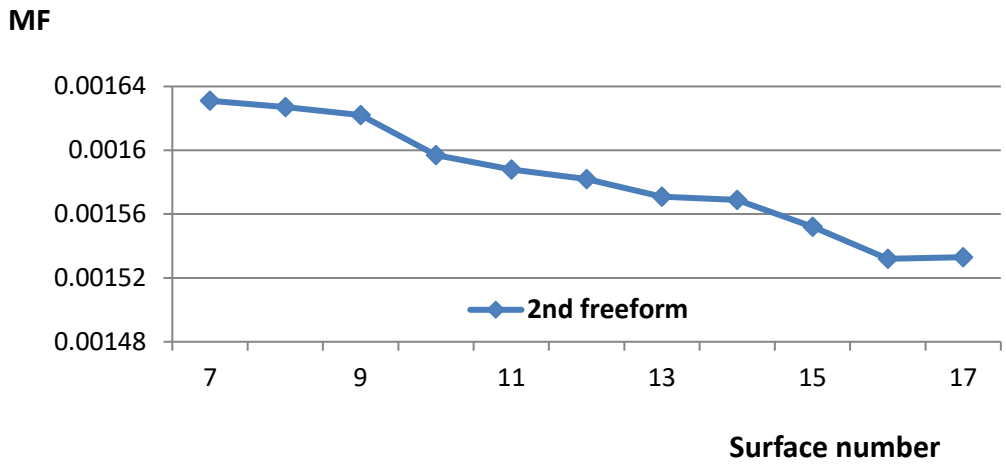



**Figure 61.** Performance of the system with the second freeform placed at different surface locations.

**Table 15**. System comparison between the start system and the freeform system. With one extra special surface, both system resolution and telecentricity are improved drastically.

| | Special surfaces | Averaged RMS spot radius | Telecentricity | Distortion |
|---|---|---|---|---|
| **Start system** | Biconic 17 | 67.09 μm | ≤7.09X10$^{-3}$ | -3.14% |
| **Freeform system** | Freeform 6,17 | 8.39μm | ≤7.07X10$^{-4}$ | -3.11% |
| **Improv. factor** | --- | 8 | 10 | 1.01 |

## 5.6 REFLECTIVE ZOOM SYSTEM

Traditional zoom systems are typically constructed with moving refractive elements, the concept of an all-reflective zoom system can be realized with the help of freeforms [70, 71]. Zoom telescope configurations discussed in this section are unobscured three-mirror systems with low f/# and large field of view (FOV). The original system shown in Figure 62 has three conic surfaces. The focal length varies from 75mm to 150mm with three moving mirrors. The system stop is placed at mirror 1. In general, mirror systems are insensitive to wavelength variation, therefore the system wavelength is set to be 1μm. System fields and f/# differ for different configurations. The f/# is in the range of 3.2 to 5.8, while the FOV varies between 2° and 3.2°.

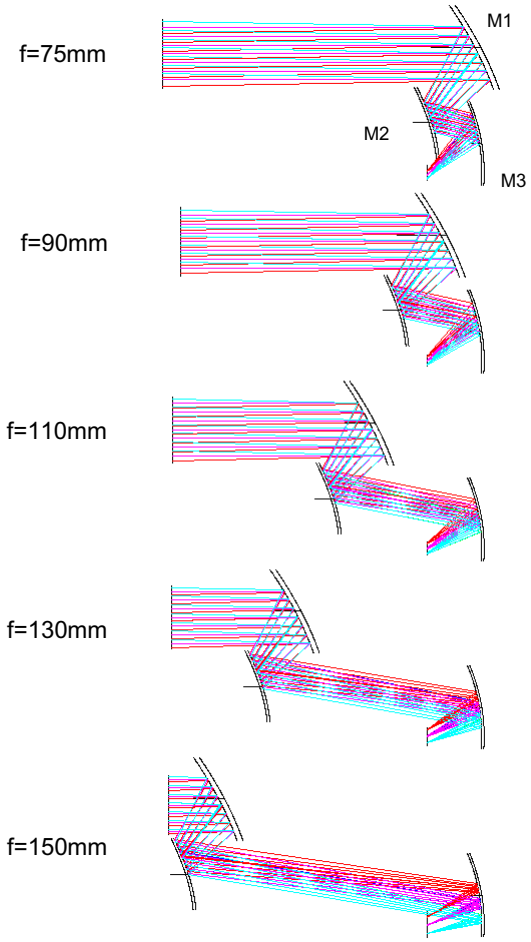



**Figure 62.** Layout of the original reflective zoom system with three conic surfaces.

The original system suffers from large field-dependent aberrations; the selection of freeform location becomes clear under the guidance of the workflow. With two freeforms placed at M2 and M3, which locate away from the stop for field-dependent-aberration correction, the layouts of the optimized system are indicated in Figure 63. The movements of the individual elements are also optimized marginally for the best results. The new system remains to be obscuration free with slightly decreased f/#. As shown in Table 16, while the FOV and focal length remain constant, the f/# of the new system decrease for configuration 3, 4 and 5. Especially in the case of configuration 5, the f/# is improved from 5.8 to 4.6. While the system distortion is controlled at a small amount for both the original system and the optimized system, the improvement in resolution is observed for all configurations with a maximum improvement factor of nearly 4.

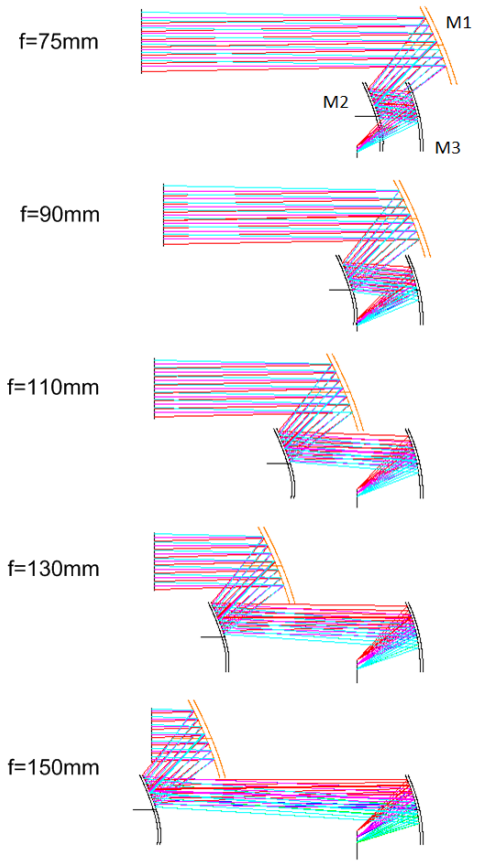



**Figure 63.** Layout of the reflective zoom system with two freeforms placed at M2 and M3.

**Table 16.** System improvement of the new system compared with the start system for all configurations. Distortion is controlled to remain at a reasonable range.

| | Focal length (mm) | FOV | Only conics | | With 2 freeforms | | Improv. factor |
|---|---|---|---|---|---|---|---|
| | | | F/# | RMS spot (μm) | F/# | RMS spot (μm) | |
| **Config.1** | 75 | -1°~1° | 3.2 | 158.9 | 3.2 | 70.2 | 2.3 |
| **Config.2** | 90 | -1.1°~1.1° | 3.4 | 99.4 | 3.4 | 36.4 | 2.7 |
| **Config.3** | 110 | -1.2°~1.2° | 3.9 | 145.1 | 3.7 | 38.6 | 3.8 |
| **Config.4** | 130 | -1.4°~1.4° | 4.7 | 55.4 | 4.1 | 30.4 | 1.8 |
| **Config.5** | 150 | -1.6°~1.6° | 5.8 | 120.7 | 4.6 | 50.4 | 2.4 |

The advantages of freeforms in system performance enhancement are demonstrated with all the given examples. The credibility of the workflow has also been validated using diverse system types. The conclusions we have made in the above work are expected to have a wider application than what has been established here.

# 6 Conclusions

In this work, an aberration analysis tool based on the Aldis theorem and implementation strategies for freeforms in multi-surface imaging systems were investigated.

In regards to the aberration analysis, the author investigated the application of the Aldis theorem based on Brewer's extension. Overcoming the limitations of existing formulas, the Aldis theorem was extended to a finite x field and to an arbitrary image plane. The original theorem is valid for surface aberration contributions of one specific ray. This limitation was partly compensated by taking into account the results of multiple ray-tracing. An extension of the original theorem to 3D systems was also achieved, in which a base ray is used as the foundation for parabasal ray tracing. The author conducted the demonstration of the newly extended Aldis theorem on two example freeform systems, for which the system sensitivities were straightforward indicated with the all-order aberration consideration.

In terms of freeform implementation strategies, two distinct cases, the one-freeform optimization and the two-freeform optimization were discussed. Impacts of the choice of the freeform normalization radius, the choice of the freeform representation orders were investigated, where the correlation was linked between the system aberration constitution and the representation orders needed from the freeform. On the other side, the selection rules of the freeform surface location were revealed with the assistance of the surface eccentricity parameter. Depending again

on the system aberration constitution, freeform surfaces were better placed closer to the stop location for field-independent aberration dominated systems, and better placed close to fields for field-dependent aberration dominated systems.

In the investigation of the two-freeform optimization, different optimization strategies were compared and it was determined that it was beneficial to add freeforms successively according to the performance requirement. The interaction between two freeforms was observed, though it is still not clear whether the interplay between freeform sags is related to the limitations of current optimization algorithm or not.

Based on the results of the two-freeform implementation investigation, an efficient workflow was proposed as guidance for freeform application. Taking the system aberration constitution into consideration, a successive implementation strategy was recommended for increasing freeform orders, as well as for increasing freeform quantity. With the preliminary manufacturability assessment of the freeform sag of a Scheimpflug system, the complex criteria was simplified and a fast evaluation for the freeform selection, taking into account the realization issues, was provided. The possibility of freeform sag control was also indicated.

In order to broaden the system selections, various system types with different system symmetries were selected as examples to demonstrate the importance of a good freeform location, as well as the system performance improvement using freeforms. The findings summarized from a small number of systems in the previous work were able to be verified on more examples. It was also demonstrated that, though a limited number of systems was investigated within the scope of this work, the conclusions are expected to have a broader application.

# Appendix

**Table 17.** Ordering and symmetry of the Zernike fringe representation up to the 10th order. The cross sign shows the symmetry information of the corresponding Zernike fringe term [35].

| Coefficient | x-symmetry | y-symmetry | x-y-symmetry | order |
|---|---|---|---|---|
| Z1 | X | X | X | 0th |
| Z2 | | X | | 2nd |
| Z3 | X | | | 2nd |
| Z4 | X | X | X | 2nd |
| Z5 | X | X | X | 4th |
| Z6 | | | | 4th |
| Z7 | | X | | 4th |
| Z8 | X | | | 4th |
| Z9 | X | X | X | 4th |
| Z10 | | X | | 6th |
| Z11 | X | | | 6th |
| Z12 | X | X | X | 6th |
| Z13 | | | | 6th |
| Z14 | | X | | 6th |
| Z15 | X | | | 6th |
| Z16 | X | X | X | 6th |
| Z17 | X | X | X | 8th |
| Z18 | | | | 8th |
| Z19 | | X | | 8th |
| Z20 | X | | | 8th |
| Z21 | X | X | X | 8th |
| Z22 | | | | 8th |
| Z23 | | X | | 8th |
| Z24 | X | | | 8th |
| Z25 | X | X | X | 8th |
| Z26 | | X | | 10th |
| Z27 | X | | | 10th |
| Z28 | X | X | X | 10th |
| Z29 | | | | 10th |
| Z30 | | X | | 10th |
| Z31 | X | | | 10th |
| Z32 | X | X | X | 10th |

| | | | | |
|---|---|---|---|---|
| Z33 | | | | $10^{th}$ |
| Z34 | | X | | $10^{th}$ |
| Z25 | X | | | $10^{th}$ |
| Z36 | X | X | X | $10^{th}$ |

**Table 18.** Ordering and symmetry of the Q-polynomial representation up to the $10^{th}$ order. The cross sign shows the symmetry information of the corresponding Q-polynomial term [35].

| Coefficient | x-symmetry | y-symmetry | x-y- symmetry | Order |
|---|---|---|---|---|
| a10 | | X | | $2^{nd}$ |
| b10 | X | | | $2^{nd}$ |
| a00 | X | X | X | $4^{th}$ |
| a11 | | X | | $4^{th}$ |
| b11 | X | | | $4^{th}$ |
| a20 | X | X | X | $4^{th}$ |
| b20 | | | | $4^{th}$ |
| a01 | X | X | X | $6^{th}$ |
| a12 | | X | | $6^{th}$ |
| b12 | X | | | $6^{th}$ |
| a21 | X | X | X | $6^{th}$ |
| b21 | | | | $6^{th}$ |
| a30 | | X | | $6^{th}$ |
| b30 | X | | | $6^{th}$ |
| a02 | X | X | X | $8^{th}$ |
| a13 | | X | | $8^{th}$ |
| b13 | X | | | $8^{th}$ |
| a22 | X | X | X | $8^{th}$ |
| b22 | | | | $8^{th}$ |
| a31 | | X | | $8^{th}$ |
| b31 | X | | | $8^{th}$ |
| a40 | X | X | X | $8^{th}$ |
| b40 | | | | $8^{th}$ |
| a03 | X | X | X | $10^{th}$ |
| a14 | | X | | $10^{th}$ |
| b14 | X | | | $10^{th}$ |
| a23 | X | X | X | $10^{th}$ |
| b23 | | | | $10^{th}$ |
| a32 | | X | | $10^{th}$ |
| b32 | X | | | $10^{th}$ |
| a41 | X | X | X | $10^{th}$ |
| b41 | | | | $10^{th}$ |
| a50 | | X | | $10^{th}$ |
| b50 | X | | | $10^{th}$ |

# References


1. M. Beier, J. Hartung, T. Peschel, C. Damm, A. Gebhardt, S. Scheiding, D. Stumpf, U. Zeitner, S. Risse, R. Eberhardt, A. Tünnermann, "Development, fabrication, and testing of an anamorphic imaging snap-together freeform telescope," Appl. Opt. 54, 3530-3542 (2015)
2. J. Stock, A. Broemel, J. Hartung, D. Ochse, H. Gross, "Description and reimplementation of real freeform surfaces," Appl. Opt.56, 391-396 (2017)
3. H. Gross, “Freeform surfaces ”, internal report (FSU Jena, 2016)
4. Arthur Cox, A System of Optical Design (FOCAL PRESS, 1964, 13-25,129-132,517-520)
5. S. H. Brewer, "Surface-contribution algorithms for analysis and optimization," J. Opt. Soc. Am. 66, 8-13 (1976)
6. S. H. Brewer, "Aberration Analysis Via Exact Surface Contributions", Proc. SPIE 0294, (1982)
7. K.P. Thompson, "Aberrations fields in tilted and decentered optical systems," Ph.D. dissertation ( University of Arizona, Tucson, Arizona,1980)
8. K. P. Thompson, "Description of the third-order optical aberrations of near-circular pupil optical systems without symmetry," J. Opt. Soc. Am. A22, 1389-1401 (2005)
9. K. P. Thompson, "Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: spherical aberration," J. Opt. Soc. Am. A 26, 1090-1100 (2009)

10. K. P. Thompson, "Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: the comatic aberrations," J. Opt. Soc. Am. A 27, 1490-1504 (2010)
11. K. P. Thompson, "Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: the astigmatic aberrations," J. Opt. Soc. Am. A 28, 821-836 (2011)
12. K. Fuerschbach, J. P. Rolland, K. P. Thompson, "A new family of optical systems employing φ-polynomial surfaces," Opt. Express 19, 21919-21928 (2011)
13. K. Fuerschbach, J. P. Rolland, K. P. Thompson, "Extending Nodal Aberration Theory to include mount-induced aberrations with application to freeform surfaces," Opt. Express 20, 20139-20155 (2012)
14. K. Fuerschbach, J. R. Rolland, K. P. Thompson, "Theory of aberration fields for general optical systems with freeform surfaces," Opt. Express 22, 26585-26606 (2014)
15. A. Yabe, "Optimal selection of aspheric surfaces in optical design," Opt. Express 13, 7233-7242 (2005)
16. A. Yabe, "Construction method of axially asymmetric lenses," Appl. Opt. 50, 3369-3374 (2011)
17. A. Yabe, "Method to allocate freeform surfaces in axially asymmetric optical systems", Proc. SPIE 8167, Optical Design and Engineering IV, 816703 (2011)
18. C. Miñano, P. Benítez, W. Lin, J. Infante, F. Muñoz, and A. Santamaría, "An application of the SMS method for imaging designs," Opt. Express 17(26), 24036–24044 (2009).
19. Y. Nie, F. Duerr and H. Thienpont, "Direct design approach to calculate a two-surface lens with an entrance pupil for application in wide field-of view imaging", Opt. Eng. 54, 015102 (2015).
20. Y. Zhong and H. Gross, "Initial system design method for nonrotationally symmetric systems based on Gaussian brackets and Nodal aberration theory", Opt. Express. 25(9) (2017).

21. S. Chang, J. Lee, S. Kim, H. Kim, W. Kim, I. Song, and Y. Park, "Linear astigmatism of confocal off-axis reflective imaging systems and its elimination," Appl. Opt.45, 484-488 (2006)
22. S. Chang, "Linear astigmatism of confocal off-axis reflective imaging systems with N-conic mirrors and its elimination," J. Opt. Soc. Am. A32, 852-859 (2015)
23. S. Mao, Y. Li, J. Jiang, S. Shen, K. Liu, M. Zheng, "Design of a hyper-numerical-aperture deep ultraviolet lithography objective with freeform surfaces," Chin. Opt. Lett. 16, 030801- (2018)
24. A. Broemel, U. Lippmann, H. Gross, "Freeform surface descriptions. Part II: Mathematical representations", Adv. Opt. Tech.6(5), 327-336 (2017)
25. A. Broemel, C. Liu, Y. Zhong, Y. Zhang, H. Gross, "Freeform surface descriptions. Part II: Application benchmark", Adv. Opt. Tech. 6 (5), 337-347 (2017)
26. M. Oleszko, R.Hambach, H. Gross, "Decomposition of the total wave aberration in generalized optical systems," J. Opt. Soc. Am. A 34, 1856-1864 (2017)
27. J. Rogers, "A comparison of anamorphic, keystone, and Zernike surface types for aberration correction." Proc. SPIE-OSA 7652, 76520B (2010)
28. R. V. Shack, K. P. Thompson, "Influence of alignment errors of a telescope system," Proc. SPIE 251, 146–153 (1980).
29. H. H. Hopkins, *The Wave Theory of Aberrations* (Oxford on Clarendon Press, Oxford, Uk,1950)
30. R. A. Buchroeder, "Tilted component optical systems," Ph.D. dissertation (University of Arizona, Tucson, Arizona, 1976)
31. J. Sasián, "Theory of sixth-order wave aberrations," Appl. Opt. 49, D69-D95 (2010)
32. A. O. Estandarte, "Contributions to the optical aberration coefficients of high order I: using Aldis Theorem and the Method of Andersen,", Revista Mexicana de Fisica **42**, No.5, 707-723 (1996)

33. B. Chen, A. M. Herkommer, "High order surface aberration contributions from phase space analysis of differential rays," Opt. Express 24, 5934-5945 (2016).
34. B. Chen, A. M. Herkommer, "Generalized Aldis theorem for calculating aberration contributions in freeform systems," Opt. Express 24, 26999-27008 (2016)
35. A. Broemel, “Development and evaluation of freeform surface descriptions,” Ph.D. dissertation (FSU Jena, 2018)
36. G. W. Forbes, "Shape specification for axially symmetric optical surfaces," Opt. Express 15, 5218-5226 (2007)
37. G. W. Forbes, "Characterizing the shape of freeform optics," Opt. Express 20, 2483-2499 (2012)
38. J. Hartung, M. Beier, S. Risse, "Novel applications based on freeform technologies", Proc. SPIE 10692, Optical Fabrication, Testing, and Metrology VI, (2018)
39. H. Gross, *Handbook of Optical Systems: Vol. 1. Fundamentals of Technical Optics* (WILEY-VCH 2005)
40. A. K. Forrest, “Oblique imaging, magnification and image intensity,” Pure Appl. Opt.**1**,687 (1999)
41. W. Richter, G. Jahn, “Some remarks on the skew collinear imaging,” Optik **111**,215 (2000)
42. A. Miks, J. Novak, P. Novak, "Analysis of imaging for laser triangulation sensors under Scheimpflug rule," Opt. Express 21, 18225-18235 (2013)
43. Y. Zhong, “Imaging with Scheimpflug setup,” master thesis report (FSU Jena, 2014)
44. H. Gross, “Alvarez plate layout ”, internal report (FSU Jena, 2016)
45. C. Spiess, “Correcting Alvarez Plates with Additional Degrees of Freefom”, internship report (FSU Jena, 2017)
46. S. S. Rege, T. S. Tkaczyk, M. R. Descour, "Application of the Alvarez-Humphrey concept to the design of a miniaturized scanning microscope," Opt. Express 12, 2574-2588 (2004)

47. P. Mouroulis, R. O. Green, B. V. Gorp, L. B. Moore, D. W. Wilson, and H. A. Bender, “Landsat swath imaging spectrometer design,” Opt. Eng. 55(1) (2016)
48. T. Eversberg and K. Vollmann, Spectroscopic Instrumentation, Springer (2015)
49. A. Romoli, F. Simonetti, L. Gambicorti, A. Z. Marchi, "New family of reflective spectrometers," Opt. Eng. 50(1) (2011)
50. W. A. Traub, "Constant-dispersion grism spectrometer for channeled spectra," J. Opt. Soc. Am. A 7, 1779-1791 (1990)
51. M. Uehara, C. Nagashima, K. Sugitani, M. Watanabe, S. Sato, T. Nagata, M. Tamura, N. Ebizuka, A. J. Pickles, K. W. Hodapp, Y. Itoh, M. Nakano, K. Ogura, "Development of the Wide Field Grism Spectrograph 2", Proc. SPIE 5492, Ground-based Instrumentation for Astronomy, (2004)
52. C. W. Chen, E. W. Gossett, “Grism (grating-prism combination),” U.S. Patent No.5,652681 (1997)
53. X. Prieto-Blanco, C. Montero-Orille, B. Couce, and R. de la Fuente, “Analytical design of an Offner imaging spectrometer,” Opt. Express14, 9156-9168 (2006)
54. M. D. Mouriz, E. L. Lago, X. Prieto-Blanco, H. González-Núñez, and Raúl de la Fuente, “Schwarzschild spectrometer,” Appl. Opt. 50, 2418-2424 (2011)
55. L. G. Cook, “Three mirror anastigmatic optical system,” U.S. Patent No. 4,265,510 (1981)
56. H. Zhu, Q. Cui, M. Piao, and C. Zhao, “Design of a dual-band MWIR/LWIR circular unobscured three-mirror optical system with Zernike polynomial surfaces,” Proc. SPIE 9272, Optical Design and Testing VI, 92720W (2014)
57. T. H. Kim, H. J. Kong, T. H. Kim, and J. S. Shin, “Design and fabrication of a 900-1700 nm hyper-spectral imaging spectrometer,” Optics Communications 283, 355-361 (2010)
58. Y. An, Q. Sun, Y. Liu, C. Li, and Z. Wang, “The design of astigmatism-free crossed Czerny-Turner spectrometer,” Optik - International Journal for Light and Electron Optics, Volume 124, Issue 16, 2539-2543 (2013)
59. R. H. El-Maksoud, M. Hillenbrand, S. Sinzinger, "Parabasal theory for plane-symmetric systems including freeform surfaces," Optical Engineering 53(3), 031303 (2013).

60. H. H. Hopkins, "Image formation by a general optical system. 1: General theory," Appl. Opt. 24, 2491-2505 (1985)
61. H. H. Hopkins, "Image formation by a general optical system. 2: Computing methods," Appl. Opt. 24, 2506-2519 (1985)
62. Y. Zhong and H. Gross, "Improvement of Scheimpflug systems with freeform surfaces," Appl. Opt. 57, 1482-1491 (2018)
63. Y. Nie, R. Mohedano, P. Benítez, J. Chaves, J. Miñano, H. Thienpont, and F. Duerr, "Multifield direct design method for ultrashort throw ratio projection optics with two tailored mirrors," Appl. Opt. 55, 3794-3800 (2016)
64. C. Liu and H. Gross, "Numerical optimization strategy for multi-lens imaging systems containing freeform surfaces," Appl. Opt. 57, 5758-5768 (2018)
65. A. Epple, T. Gruner, W. Singer, "Optical system for ultraviolet light," US. Patent 2006/0119750 (2006)
66. C. Liu, C. Straif, T.Flügel-Paul, U. D. Zeitner, H. Gross, "Optical design and tolerancing of a hyperspectral imaging spectrometer", Proc. SPIE 9947, Current Developments in Lens Design and Optical Engineering XVII, 994703 (2016)
67. C. Liu, C. Straif, T. Flügel-Paul, U. Zeitner, H. Gross, "Comparison of hyperspectral imaging spectrometer designs and the improvement of system performance with freeform surfaces," Appl. Opt.56, 6894-6901 (2017)
68. A. Grewe, M. Hillenbrand, S. Sinzinger, "Aberration analysis of optimized Alvarez–Lohmann lenses," Appl. Opt. 53, 7498-7506 (2014)
69. C. Liu, H. Gross, "Investigation of optimization strategy and freeform location on multi-surface systems", Proc. SPIE 10690, Optical Design and Engineering VII, 106901G (2018)
70. R. B. Johnson, "Unobscured reflective zoom systems," Proc. SPIE 2539, Zoom Lenses, (1995)
71. X. Wang, J. Chang, B. Shen, W. Zha, Y. Niu, K. Zhang, "Design of an all-reflective active zoom system," Optik, 1022-1026 (2016)

# List of Publications

I. C. Liu, C. Straif, T. Flügel-Paul, U. Zeitner, H. Gross, "Comparison of hyperspectral imaging spectrometer designs and the improvement of system performance with freeform surfaces," Appl. Opt.56, 6894-6901 (2017)

II. C. Liu and H. Gross, "Numerical optimization strategy for multi-lens imaging systems containing freeform surfaces," Appl. Opt. 57, 5758-5768 (2018)

III. C. Liu, C. Straif, T.Flügel-Paul, U. D. Zeitner, H. Gross, "Optical design and tolerancing of a hyperspectral imaging spectrometer", Proc. SPIE 9947, Current Developments in Lens Design and Optical Engineering XVII, 994703 (27 September 2016)

IV. C. Liu, H. Gross, "Investigation of optimization strategy and freeform location on multi-surface systems", Proc. SPIE 10690, Optical Design and Engineering VII, 106901G (5 June 2018)

V. A. Broemel, C. Liu, Y. Zhong, Y. Zhang, H. Gross, "Freeform surface descriptions. Part II: Application benchmark", Adv. Opt. Tech. 6 (5), 337-347 (2017)

# List of Symbols

$\vec{H}$ field position in the image field

$\vec{\rho}$ pupil position

$W_{klm}$ wave aberration expansion coefficient

$\vec{\sigma}_j$ decentration of the center of the aberration field with respect to the unperturbed center

$\vec{H}_{Aj}$ effective aberration field height of surface j

$(l_s, m_s, n_s)$ optical direction cosines of ray after refraction at surface s

$(X_s, Y_s, Z_s)$ real ray intersection Cartesian coordinates at surface s

$\beta_s$ paraxial marginal ray height at the tangential plane at surface s

$\mu_s$ refractive index following surface s

$b_s$ paraxial marginal ray angle after refraction at the tangential plane of surface s times $\mu_s$

$Y_p$ Gaussian image height at the tangential plane

$U_s$ aberration contributions in x of individual surface s

$V_s$ aberration contributions in y of individual surface s

$\Delta Y$ total transverse aberration in y for one ray

$t_{s-1}$ distance between surface s-1 and surface s

$\beta x_s$ paraxial marginal ray height at the sagittal plane of surface s

$bx_s$ paraxial marginal ray angle after refraction at the sagittal plane of surface s times $\mu_s$

$X_p'$ paraxial image height at the sagittal plane of the shifted image plane.

$\Delta x_n$ chief ray height different at the sagittal plane due to the image plane shift

$A(\bar{x}, \bar{y})$ boundary function

$P(x, y)$ projection factor

F $(\bar{x}, \bar{y})$ freeform deformation

$\bar{x}, \bar{y}$ normalized Cartesian coordinates

$W(x, y)$ wavefront aberration

$Z_n^m(\bar{r}, \varphi)$ Zernike term

$c$ surface curvature

$Q_n^m(v)$ polynomial of order n in v.

$\theta_s$ angle formed between the base ray and the z axis after surface s

$\overline{\beta x_s}$ parabasal marginal ray height at the sagittal plane of surface s

$\overline{bx_s}$ parabasal marginal ray angle after refraction at the sagittal plane of surface s times $\mu_s$

$\overline{\beta_s}$ parabasal marginal ray height at the tangential plane of surface s

$\overline{b_s}$ parabasal marginal ray angle after refraction at the tangential plane of surface s times $\mu_s$

$X_s^b$ x intersection height of the base ray at surface s.

$P, Q, P', Q'$ correction terms for parabasal ray tracing

$m$ magnification

$k$ normalization factor

$\chi$ surface eccentricity

$h_{CR}$ absolute chief ray height

$h_{MR}$ absolute marginal ray height

$M_{ij}$ system performance for freeform choice i and j, with i being the first freeform, j being the second freeform

# Ehrenwörtlichen Erklärung

Hiermit erkläre ich ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig, ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als die angegebenen Hilfsmittel und Literatur angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quellen gekennzeichnet.

Bei der Auswahl und Auswertung folgender Materialien haben mir nachstehenden Personen in der jeweiligen Weise entgeltlich/unentgeltlich geholfen:

- Herbert Gross, Betreuer
- Das Freiformdarstellung „Q-Polynome“ wurde gemeinsam von Uwe Lippmann und Anika Brömel in Zemax implementiert.
- Die Abbildungen 58 und 60 wurden mit dem von Shenlei Zhu verfassten Zemax-Makro erstellt, der diese Arbeit unter der Aufsicht von Dr. Yi Zhong durchführte.
- Die Makros zur Berechnung von Aldis Theorem benutzten ein Toolkit, welches von Joseph M. Howard (2017, NASA) geschrieben wurde und unter “http://opensource.gsfc.nasa.gov/projects/Matlab_Zemax/index.php” veröffentlicht ist.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs-bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar und mittelbar geldwerte

Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.
Weder diese Arbeit noch eine andere wurde bisher im In- oder Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.
Die geltende Promotionsordnung der Physikalisch-Astronomischen Fakultät ist mir bekannt. Ich versichere ehrenwörtlich, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Jena, 21.01.2019 Chang Liu

# Acknowledgements

I would like to thank Professor Herbert Gross for his guidance and coaching regarding optical design through my complete study, as well as his constant encouragement and support during my research years in Jena. It is his vision that predicts the potential of this work and sees through it from the very beginning. This complete thesis would not have been possible without his fresh ideas and constant motivation.

Many thanks go to my colleagues, Yi Zhong, Anika Brömel, Mateusz Olesko, Christoph Bösel, Britta Satzer, Johannes Stock, Uwe Lippmann and Johannes Hartung for the countless fruitful discussions regarding freeforms, which have helped me tremendously. I want to thank my colleague Norman Girma Worku for his attempt in helping me to understand 3D ray tracing. My appreciation goes to Sebastian Merx for his implementation of Yabe's first method as a Zemax macro. I also would like to thank Ewa Kasperkiewicz, the knowledge I obtained regarding Scheimpflug system during my supervision of her internship work has enlightened this Ph.D. work. Thanks also go to all my colleagues from the optical system design group for their kindness and constant emotional support.

Lastly, my deepest gratitude goes to my parents, who gave me the most valuable gift of life 27 years ago. None of these would become the reality without their constant support, unending patience and selfless love.